Folding and Aggregation of IgGs - Influence of Mutations and Expression Systems

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Jonas V. Schaefer

aus

Deutschland

Promotionskomitee

Prof. Dr. Andreas Plückthun (Vorsitz)

Prof. Dr. Inger Sandlie

Prof. Dr. Peter Sonderegger

Prof. Dr. Ari Helenius

Zürich, 2012

Die vorliegende Arbeit wurde von der Mathematisch-naturwissenschaftlichen Fakultät der Universität Zürich im Sommersemester 2012 als Dissertation angenommen.
Promotionskomitee: Prof. Dr. Andreas Plückthun (Vorsitz und Leitung der Dissertation), Prof. Dr. Inger Sandlie, Prof. Dr. Peter Sonderegger, Prof. Dr. Ari Helenius

To my family

"The most exciting phrase to hear in science, the one that heralds the most discoveries, is not 'Eureka!' (I found it!) but 'That's funny...' " ~ Isaac Asimov

Erklärung

Diese Dissertation wurde selbstständig, ohne unerlaubte Hilfe im Sinne von §3 und §5 der Promotionsverordnung vom 08. Juli 2002 angefertigt. Bei der Abfassung der Dissertation wurden keine anderen als die darin angegebenen Hilfsmittel benützt.

Zürich, Januar 2012

Jonas V. Schaefer

Acknowledgments

This thesis and my research at the University of Zurich would not have been possible without the kind assistance, practical contributions and great support of the following individuals and organizations:

My foremost thanks go to my supervisor, Prof. Dr. Andreas Plückthun, for his encouragement, guidance and continuous support during my PhD studies as well as the freedom to pursue my own ideas and thoughts. Our discussions and his valuable feedback enabled me to accomplish hypothesis-driven research and I am thankful for the ability to gain insight into many different aspects of antibody engineering and various lab techniques. I also want to extend my gratitude to the members of my thesis committee, Prof. Dr. Inger Sandlie (Centre for Immune Regulation, University of Oslo, Norway), Prof. Dr. Ari Helenius (Institute of Biochemistry, ETH Zurich, Switzerland) and Prof. Dr. Peter Sonderegger (Department of Biochemistry, University of Zurich, Switzerland), for their helpful comments and valuable advice during the entire course of this study.

I would also like to gratefully acknowledge the support of my colleagues in the Plückthun laboratory who helped me in my research through many fruitful discussions and inspiring suggestions. Especially, many thanks to the M74 crew at various stages (Fabio, Patricia and Matthias) for such a friendly and relaxed working atmosphere and all the fun we had in our office. I owe special thanks to Dr. Peter Lindner, who - next to being a great lab manager - aided me in numerous aspects of my research through his thoughtful assistance and by critically reading various manuscripts. I further thank Dr. Ykelien Boersma for also critically reading some of my manuscripts, for valuable suggestions and for her friendship including all the yummy deserts. I am also grateful to the following individual of the Plückthun laboratory: Anja, Petra, Annemarie, Christian, Birgit, Erik, Niki, Myriam, Patricia, Gabriela and all the others not explicitly mentioned current and past colleagues of the Plückthun laboratory, in particular all those who have provided any advice or technical assistance during my thesis. Next to leading to a motivating and enjoyable atmosphere, many of them were part of various social lab activities from city trips, weekend hikings, BBQs and movie nights. I especially appreciate your help and amazing support in the months after my sledging accident – thanks guys!

In addition I´d like to thank several people from academic and industrial partners that were involved at various stages of this project. For their help with DSC, DSF and MALS analyses I am thankful to Dr. Ilian Jelezarov, Dr. Paolo Cinelli and Dr. Birgit Dreier from the University of Zurich. I want to further thank Dr. Peter Gimeson (GE Healthcare, Sweden), Dr. Thomas Müller-Späth (ChromaCon, Switzerland) and Dr. Daniel Weinfurtner (MorphoSys, Germany) for their generous willingness to support me in the DSC, CIEX and Capillary electrophoresis experiments. I would like to acknowledge Dr. Manfred Heller (University of Bern, Switzerland) for his active help in MS analyses. I am further grateful to the personnel of the Functional Genomics Center at University of Zurich for their help with MS analyses, protein sequencing and RP-HPLC. My thanks are also due to Dr. Stefan Duhr (NanoTemper, Germany) for the chance to determine affinities by microscale thermophoresis. Moreover I am indebted to the P-CUBE program and the Welcome Trust Center (Oxford, UK), especially Dr. Yuguang Zhao and Margaret Jones, for their help and support during my stay in Oxford.

I highly appreciate the great help of all the people who supported my work in terms of administration, infrastructure and materials. Thanks to the entire staff of the Department of Biochemistry, especially Petra Vogt, the IT-Team (Steve Rast and Dr. Stephan Klauser) and the workshop crew (Adrian Schmid und Sascha Weidner). Finally, I´d like to thank the Molecular Life Science PhD Program of the Life Science Zurich Graduate School for giving me the opportunity to be part of the PhD program, allowing me to get to know many other PhD students from various fields and to exchange our ideas and thoughts at various retreats and events.

Furthermore, I am very grateful for the financial support of a Kekulé-Scholarship from the 'German Chemical Industry Association' during the first two years of my PhD studies. I also want to thank the University of Zurich, the Life Science Zurich Graduate School, the Hartmann Müller Foundation as well as the Swiss Society of Biochemistry for providing several travel grants for the attendance of various conferences in Europe and in the United States.

Most importantly, I would like to thank my family for the support they provided throughout my entire life and, in particular, my stays abroad. They have always been there for me when I needed them and encouraged me continuously through all these years. I am further extremely thankful for the love and support of my girlfriend Alexandra who helped me to stay balanced and motivated once labwork was getting tough. Thanks for accepting the numerous times I left you behind going to the next conference or meeting and sorry for regularly messing up our vacation plans.

Table of Content

1. Summary

Recombinant antibodies and their derivatives are widely used molecules in a broad range of applications, from basic research to medical therapy. Although immunoglobulins (IgGs) generally belong to the more stable natural proteins, by far not all of them show the characteristics required for these tasks. Their biophysical properties can differ greatly from very stable to highly aggregation-prone. Aggregation is a serious concern for therapeutic antibodies, since it can lead to reduced bioactivity, cause severe side effects and increase the risk of immunogenicity. Therefore, the understanding and the engineering of antibody degradation and aggregation mechanisms is crucial for the development of potent pharmaceutical products. In addition, developing methods to improve the expression level and overall stability of IgG molecules is of great importance, as the high cost of their production is still a major drawback of these molecules.

Previous work from our laboratory consolidates the combination of structure-based analysis with family-consensus alignments as a powerful method for improving the properties of immunoglobulin variable domains. Particular germline residues in the framework of these domains were shown to be responsible for poor biophysical properties of some of the IgG subclasses. The exchange of these amino acids in the V_H domain greatly influenced both the stability as well as the expression level of scFv fragments produced in *Escherichia coli*. However, nowadays almost all approved therapeutic antibodies are still in the IgG format produced in mammalian cells. Therefore, the main focus of this thesis was to investigate whether these effects were unique to the scFv format and the prokaryotic expression system or whether they were applicable to other antibody formats as well. Therefore, antibody-formats with more than the variable domain (Fab fragment and full-length IgGs) were designed and different expression hosts (*E. coli*, *Pichia pastoris* and mammalian cells) selected as production systems. As so far very little is known about the extent to which results derived from one antibody format (like scFv) can be applied to other formats (like the full-length IgG), this project showed great promise for gaining further insights into the influence of intrinsic domain interactions. In addition, it was intended to elucidate to which extent the choice of the expression system with its respective protein folding and quality control machineries influence the expressed proteins.

In the first part of this thesis, the transferability of beneficial mutations to other antibody formats came to the fore. Analyses clearly indicated that - although increased expression levels were only detectable for periplasmatic expression in the prokaryotic system - advanced stabilities both with respect to thermal and denaturant-induced unfolding could be found in formats other than the scFv, independent of the expression system. Interestingly, the mutations also influenced the structural integrity of the engineered full-length IgG. An increased homogeneity was not restricted to the V_H domain but even affected the structural integrity of the adjacent C_H1 domain with its intrachain disulfide bond. These results confirm the potential of structure-based protein engineering for applications involving full-length IgGs and the transferability of improvements implemented in smaller antibody fragments.

The comparison of the different expression hosts led to an additional important finding. The analysis of IgGs of identical amino acid sequence that were produced either in mammalian cells (HEK293) or in the yeast *P. pastoris* detected dramatic differences in some of their biophysical characteristics. While the thermodynamic stabilities of the constructs were quite similar, they unexpectedly revealed very different aggregation susceptibilities: Antibodies expressed in mammalian cells aggregated at ~75°C whereas the analogous constructs from *P. pastoris* did not show any detectable aggregation. This phenomenon was found to be mainly caused by two factors. As a first cause, differences in the glycosylation pattern attached to the C_H2 domain of the IgGs were identified. While the mannose-rich glycan found in the *Pichia* IgGs slightly destabilized the molecules with respect to thermal and denaturant-induced unfolding, it strongly inhibited their aggregation propensity compared to the complex mammalian glycan.

Secondly, the increased tolerance towards aggregation of the *Pichia*-derived antibodies was found to be caused by residues left behind from the α-factor pre-pro sequence at the N-termini of both the heavy and light chain due to imperfect processing. This additional EAEA-tetrapeptide proved to increase the temperature of the onset of aggregation considerably and reduced aggregate formation after extended incubation at elevated temperatures. The experimental attachment of these residues to IgGs produced in the mammalian cell culture confirmed their beneficial effect on aggregation resistance. While the extension with control tetrapeptides even increased the aggregation propensity of mammalian IgGs, the addition of EAEA to the light chain - but not to the heavy chain - resulted in an apparent increase in the onset temperature of aggregation by approximately 10°C. Data derived from analyses performed at various pH support the hypothesis that the positive impact of the EAEA-tetrapeptide most likely is based on its contribution towards the local charge distribution of the antibody molecule.

To warrant an extensive comparison of IgG molecules, great efforts were made to enable secretion of IgGs with native N-termini in the yeast system. Systematic engineering of the precursor proteins and the processing site was performed and showed that the heavy chain did not require the pro-region for efficient transport and folding in *Pichia*. These results led to the definition of optimal sequence composition for either aggregation-resistant or correctly processed IgGs, respectively.

Taken together, the second part of this thesis investigated the impact of certain sequences on the aggregation properties of IgGs and showed that their aggregation susceptibilities could be engineered by modifying their charge distributions at certain locations. These results will hopefully improve the engineering of aggregation-resistant IgGs and further be useful for the efficient production of full-length IgGs in *P. pastoris*. In addition, the data shed new light on potential biophysical effects of tag sequences in general. As could be shown, the addition of only a few amino acids to the N-terminus of the IgG, a protein of several hundred residues, had dramatic effects on the molecule's biophysical properties *in vitro*. In a broader view, these findings have important implications for the common practice of adding different tags to proteins of interest, assuming that few additional amino acids will not have major impacts on these molecules.

As a last part of this thesis, various practical aspects of antibody production were addressed. Developing and enhancing previous techniques and methods, improved instructions could be established allowing the direct application of this part to practical challenges. By designing a new set of primer sequences for the amplification of V_H and V_L genes, the faithful cloning of the variable region genes could be promoted, preserving as much sequence identity as possible while avoiding the generation of non-natural residue combinations which could result in sequences problematic for folding and stability. In an additional approach, methods to create multivalent molecules were elucidated. These designed "miniantibodies" mimic the geometry of IgG molecules and thus increase their functional affinity (avidity) to a corresponding multimeric antigen structure. At the same time, the used components can conveniently be produced in bacterial systems. Since, furthermore, two different specificities can be combined within the same miniantibody, these molecules allow numerous applications in biotechnology, diagnostics and potentially therapy. Allowing the crosslinking of two targets, this approach is of special interest as bispecific binders have already proven to be of great therapeutical benefit. Finally, aspects of antibody fragment production in *E. coli* were addressed. Due to their conserved intradomain disulfide bonds, antibody fragments need to be secreted to the oxidizing compartment of the periplasm for correct folding. As the periplasmic folding, however, is the yield-limiting step, co-expression of periplasmic factors is thought to improve the yield of correctly folded antibodies. This problem of soluble expression of scFvs was tackled by creating a modular system that allows a flexible co-expression of many factors with virtually any antibody expression vector.

Summarized, this thesis deals with various aspects of antibodies and challenges of their application, highlighting ways how to engineer and improve their properties for the next generation of IgGs.

Zusammenfassung

Rekombinant-hergestellte Antikörper und ihre Derivate sind für viele Aufgabengebiete - angefangen von der Grundlagenforschung bis hin zu klinischen Anwendungen - von besonderem Interesse. Immunglobuline (IgGs) gehören zu den eher stabilen natürlichen Proteinen, dennoch weisen nur wenige dieser Moleküle die für ihre zahlreichen Anwendungen benötigten Merkmale auf. Besonders ihre biophysikalischen Eigenschaften können hierbei deutlich variieren, angefangen von extrem stabil bis hin zu höchst aggregationsanfällig. Aggregation ist für therapeutische Antikörper von großer Bedeutung, da es sowohl zu einer reduzierten Bioaktivität führen als auch das Risiko der Immunogenität deutlich erhöhen kann. Daher ist es für die Entwicklung von erfolgreichen pharmazeutischen Produkten entscheidend, die Abbau- und Aggregationsmechanismen von Antikörpern besser zu verstehen und zu beeinflussen. Zusätzlich sollten Methoden zur Verbesserung der Expressionsmengen sowie der generellen Stabilität von IgG Antikörpern entwickelt werden, da deren hohe Produktionskosten immer noch einen bedeutenden Nachteil darstellen.

Frühere Arbeiten unserer Arbeitsgruppe haben gezeigt, dass die Kombination von strukturbasierten Analysen mit Vergleichen der Konsensus-Sequenzen der IgG Familien eine leistungsfähige Methode zur Eigenschaftsverbesserung der variablen Immunglobulin-Domänen darstellt. Hierbei konnte nachgewiesen werden, dass spezielle Positionen innerhalb der in der Keimbahn codierten konstanten Strukturen für die mangelhaften Eigenschaften mancher IgG Familien verantwortlich sind. Der Austausch dieser Aminosäuren innerhalb der V_H Domäne verbesserte tatsächlich sowohl die Expressionsmenge als auch die Stabilität von scFv-Fragmenten, welche im Bakterium *Escherichia coli* hergestellt wurden. Da heutzutage jedoch fast alle therapeutisch zugelassenen Antikörper noch immer als ganze Antikörper in eukaryotischen Zellen hergestellt werden, steht im Mittelpunkt dieser Doktorarbeit die Frage, ob die eben erwähnten Effekte einzig und allein für das scFv-Format und prokaryotische Expressionssysteme gelten oder ob sie auch auf andere Antikörperformen übertragen werden können. Um dies zu untersuchen, wurden verschiedene Antikörperformate (Fab-Fragmente und ganze IgGs) ausgewählt und in verschiedenen Expressionssystemen (*E. coli*, *Pichia pastoris* sowie Säugerzellen) hergestellt. Da bislang wenig darüber bekannt ist, in welchem Ausmaß die von einem Antikörperformat (z.B. scFv-Fragmente) gewonnenen Ergebnisse auf ein anderes Format (z.B. ganze IgGs) übertragen werden können, sollte dieses Projekt weitere Erkenntnisse über den Einfluss der Wechselbeziehungen der intrinsischen Domänen liefern. Darüber hinaus galt es zu analysieren, inwiefern die verschiedenen Expressionssysteme mit den dazugehörigen Proteinfaltungs- und Qualitätskontroll-Mechanismen die Eigenschaften der hergestellten Proteine beeinflussen.

Im ersten Teil dieser Doktorarbeit stand die Frage nach der Übertragbarkeit der vorteilhaften Mutationen auf andere Antikörperformate im Vordergrund. Diese konnte positiv beantwortet werden. Verbesserte Expressionsmengen wurden zwar nur für die periplasmatische Expression im prokaryontischen System erzielt - jedoch konnte klar gezeigt werden, dass eine erhöhte Stabilität bezüglich der thermischen sowie der durch Denaturierungsmittel ausgelösten Entfaltung auch in sich von scFv unterscheidenden Formaten erreicht wird. Dies scheint unabhängig vom verwendeten Expressionssystem der Fall zu sein. Interessanterweise veränderten die Mutationen auch die

strukturelle Integrität und die Homogenität der veränderten IgG Antikörper. Diese verbesserte Homogenität betraf nicht nur die V_H Domäne, sondern beeinflusste zusätzlich auch die strukturelle Integrität der benachbarten C_H1 Domäne mit ihrer Disulfidbrücke zwischen leichter und schwerer Kette. Zusammengenommen bestätigen diese Ergebnisse das große Potenzial des strukturbasierten Antikörper-Engineerings und die Übertragbarkeit von Erkenntnissen, welche im Zusammenhang mit kleineren Antikörper-Fragmenten gewonnen wurden.

Der Vergleich der verschiedenen Expressionssysteme führte zu einer weiteren wichtigen Erkenntnis. Bei der Analyse von IgG Molekülen - welche zwar die identische Aminosäuren-Sequenz besaßen, jedoch entweder in Säugerzellen (HEK293) oder von der Hefe *P. pastoris* produziert wurden – zeigten sich erstaunliche Unterschiede in einigen biophysikalischen Eigenschaften. Während die thermodynamische Stabilität aller Konstrukte recht ähnlich war, konnten überraschenderweise doch sehr unterschiedliche Aggregationsanfälligkeiten nachgewiesen werden. Antikörper, welche in Säugerzellen hergestellt wurden, aggregierten bei ~75°C, während die entsprechenden Konstrukte aus dem Hefesystem keinerlei nachweisbare Aggregation aufwiesen. Bei den durchgeführten Untersuchungen stellte sich heraus, dass dieses Phänomen auf zwei unterschiedlichen Ursachen basiert. Als erste konnten Unterschiede bei der Glykosylierung der C_H2 Domänen identifiziert werden. Während die Zuckerstrukturen der Hefe-Antikörper zahlreiche Mannose-Einheiten aufwiesen und das Molekül geringfügig destabilisierten, verminderten sie dennoch die Aggregationsanfälligkeiten deutlich im Vergleich zu den entsprechenden Säugerzuckern.

Als zweite Ursache der vergrößerten Aggregationstoleranz der Hefe-Antikörper konnten Aminosäuren ausgemacht werden, welche von der verwendeten α-factor Pre-Pro-Sequenz herrührten und durch ein unvollständiges Prozessieren der schweren und leichten Kette jeweils an deren N-Termini zurückblieben. Dieses zusätzliche Tetrapeptid mit der Sequenz "EAEA" war für die erhöhte Starttemperatur der Aggregation verantwortlich und reduzierte darüber hinaus auch die Menge und Größe von Aggregaten nach längerer Lagerung bei erhöhten Temperaturen. Das Anhängen dieser Reste an in Säugerzellen produzierte IgG Antikörper bestätigte deren positiven Beitrag zur Verminderung der Aggregationsanfälligkeit. Während Kontrollpeptide die Resistenz in Säuger-IgGs sogar noch reduzierten, erhöhte das zusätzliche EAEA-Peptid die messbare Starttemperatur der Aggregation um rund 10°C - jedoch nur dann, wenn es an die leichte, nicht aber an die schwere Kette gehängt wurde. Aggregationsanalysen bei verschiedenen pH Werten deuten darauf hin, dass der Vorteil des EAEA-Peptids auf dessen Beitrag zur lokalen Ladungsverteilung innerhalb des Antikörpermoleküls zurückzuführen ist.

Um die Antikörper aus den verschiedenen Expressionssystemen direkt miteinander vergleichen zu können, wurden auf mehrfache Weise versucht, die Herstellung von IgG Molekülen mit ursprünglichen N-Termini im Hefesystem zu ermöglichen. Systematisches Engineering der Vorläufer-Moleküle und der benötigten Schnittstellen der beteiligten Enzyme ergab, dass die schwere Kette die verwendete Pro-Region nicht zum effizienten Transport und zur erfolgreichen Faltung benötigt. Diese Ergebnisse erlaubten die Festlegung von optimalen Sequenzzusammensetzungen für aggregationsresistente oder alternativ korrekt-prozessierte IgG Antikörper.

Durch die Untersuchung des Einflusses bestimmter Sequenzen auf das Aggregationsverhalten von IgG Antikörpern konnte gezeigt werden, dass die Aggregationsanfälligkeit durch das Verändern der Ladungsverteilungen innerhalb des Moleküls reduziert werden konnte. Diese Erkenntnisse werden, so ist zu hoffen, sowohl das Engineering von aggregations-resistenten IgGs ermöglichen als auch die effiziente Herstellung von ganzer IgGs in *P. pastoris* weiter vorantreiben. Gleichzeitig zeigen diese Resultate potenzielle Auswirkungen verschiedener Protein-Tags auf deren biophysikalische Eigenschaften. Denn schon wenige zusätzliche Aminosäuren am N-Terminus von IgG Molekülen - welche immerhin aus mehreren hundert Resten bestehen - haben einen dramatischen Einfluss auf deren *in vitro* Eigenschaften. Übergreifend weist dies auf eine bisher unterschätze Problematik hin, wenn heutzutage generell die verschiedensten Tags an Proteine gehängt werden – immer in der Annahme, dass die zusätzlichen Aminosäuren keinen großen Einfluss auf das zu untersuchende Protein haben.

Schließlich befasst sich diese Doktorarbeit noch mit verschiedenen praktischen Aspekten der Antikörperherstellung. Vorhandene Techniken und Methoden wurden weiterentwickelt, um verbesserte Anleitungen zur Verfügung stellen zu können. So konnte durch das Generieren eines neuen Primer-Sets für die Vervielfältigung der V_H und V_L Gene das Klonieren dieser variablen Domänen verlässlicher gestaltet werden. Hierbei wurde eine größtmögliche Sequenzidentität erhalten und nicht-natürliche Aminosäure-Kombinationen verhindert, welche zu Sequenzen mit problematischen Faltungs- und Stabilitätseigenschaften führen könnten. In einem weiteren Ansatz wurden Methoden zur Herstellung multivalenter Moleküle entworfen und verfeinert. Die in Kapitel 3.4 vorgestellten "Miniantikörper" ahmen den Aufbau von IgG Molekülen nach und erlauben daher eine verbesserte funktionelle Affinität (Avidität) zu multimeren Antigenen. Als großer Vorteil erweist es sich, dass die verschiedenen Komponenten bequem in prokaryotischen Systemen hergestellt werden können. Da hierbei auch zwei verschiedene Spezifitäten innerhalb desselben Miniantikörpers vereint werden können, ergeben sich für sie verschiedene technologische, diagnostische und sogar therapeutische Anwendungen. Durch die Möglichkeit, durch Miniantikörper zwei Ziele wie z.B. Zellen zu verlinken, ist dieser Ansatz von besonderem Interesse - angesichts des nachgewiesenen therapeutischen Nutzens von bispezifischen Bindern. Abschließend wurden Aspekte der Produktion von Antikörper-Fragmenten in *E. coli* behandelt. Wegen der Disulfidbrücken innerhalb ihrer Domänen müssen Antikörper-Fragmente in das oxidierende Milieu des Periplasma sekretiert werden. Da das dort stattfindende Falten jedoch in der Regel der (mengen-) limitierende Schritt ist, könnte die Koexpression von periplasmatischen Faktoren die Anzahl an korrekt gefalteten Antikörper erhöhen. Um diese Herausforderung anzugehen, wurde ein modulares System entwickelt, welches die flexible Koexpression zahlreicher Faktoren mit fast allen erdenklichen Expressionsvektoren für die Antikörperherstellung ermöglicht.

Zusammengefasst beschäftigt sich diese Dissertation sowohl mit verschiedenen Aspekten von Antikörpern als auch mit einigen Herausforderungen ihrer Anwendung und zeigt Wege auf, die Eigenschaften der nächsten IgG Generation weiterzuentwickeln und zu verbessern.

2. Introduction

Antibodies, also called immunoglobulins (Ig), are an important component of the vertebrate immune system. As part of the humoral immune response, these molecules recognize many pathogens and thereby protect vertebrates against infections of e.g. viral and bacterial origin. In addition, antibodies and their derivatives are used in an increasing number of applications from basic research to biomedical therapy and thus represent one of the most important classes of molecules in the biotechnological and pharmaceutical industry. The main reason for this wide range of applications is the ability of antibodies to selectively recognize even small quantities of many kinds of substances, from small organic components to large surface structures of microbes or mammalian cells. Moreover, antibodies can bind their targets with both extraordinary specificity as well as high affinity, making them ideal tools for a broad variety of applications. Currently, 30 therapeutic monoclonal antibodies (mAbs) have been approved for therapeutic use by the US Food and Drug Administration (FDA) and several hundred antibody-derivatives are in clinical development, making up over a third of the proteins currently undergoing clinical testing in the USA.[1; 2] Most of the therapeutic antibodies are used in the areas of oncology or the treatment of arthritis and immuno-inflammatory disorders. In cancer therapy, for example, mAbs successfully support the standard treatments of surgery and radiation.[3] In addition to their pharmaceutical relevance, antibodies are also employed in a wide range of analytical techniques like immunoblotting, flow cytometry, immunohistochemistry or enzyme-linked immune-sorbent assays (ELISA), just to name a few examples. Furthermore, antibodies are employed in specific protein purification by affinity chromatography.

Structure and function of antibodies

Immunoglobulin G (IgG), the immunoglobulin most frequently found in human serum, is a heterotetramer composed of two identical light (L) chains and two identical heavy (H) chains, having a molecular weight of ~25 kDa and ~50 kDa per chain, respectively. There are two variants of human light chains present in human IgGs, the more frequent kappa-(κ-)chain and the less frequent lambda-(λ-)chain. The two heavy chains of the IgG molecule are held together by noncovalent forces such as hydrophobic interactions or van-der-Waals forces and additionally linked by two interchain disulfide bonds (in the case of IgG1) within the hinge region (Fig. 2.1). Each of these heavy chains is further connected with one light chain by hydrophobic interactions and an additional disulfide bridge. This composition leads to the well-known Y-shaped structure of IgGs with two identical antigen-binding sites, one in each of its arms (Fab fragments). These binding sites are formed by the complementarity determining regions (CDRs) - loops which are highly diverse in length and sequence. These six hypervariable regions - three each within the N-terminal variable (V) regions of both the light and heavy chain, respectively - generate the unique surface that specifically recognizes and binds the corresponding antigen.[4; 5] The CDRs themselves are flanked by four framework regions that guarantee the stability of the variable domains and the orientation of the CDRs, all together forming the so-called paratope which specifically binds to the epitope on the antigen. The remaining IgG framework is highly conserved both in its sequence as well as in the structure of its constant regions that mediate the

biological activities characteristic for antibodies. These include the ability to initiate host immune responses such as antibody-dependent cellular cytotoxicity (ADCC) or complement-dependent cytotoxicity (CDC).[6] All these effects are mediated by the crystallizable fragment of the IgG (Fc part) consisting of the second and third constant domains of the heavy chain (C_H2 and C_H3). In addition, a complex N-linked glycosylation present at Asn297 within the C_H2 domain is required for the interaction of the IgG with some of the activating Fcγ receptors. These receptors, and among them especially FcγRIIIa, are responsible for triggering ADCC by recruiting immune cells that lead to phagocytosis or lysis of the targeted cells. In CDC, antibodies activate the complement cascade at the cell surface, eventually killing the targeted cells.

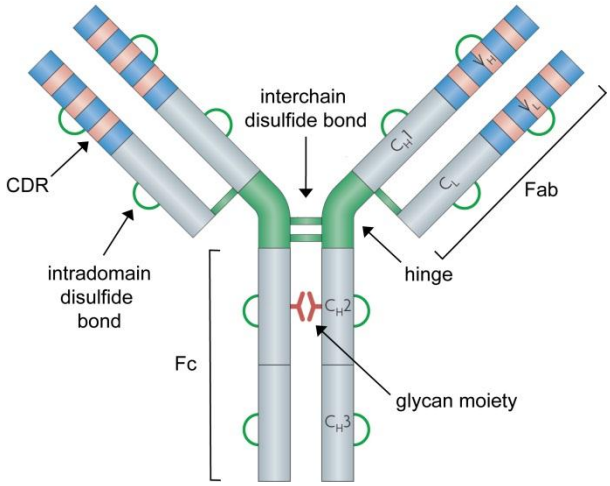


Fig. 2.1: Schematic representation of immunoglobulin G. The IgG molecule is comprised of four polypeptides, two identical heavy (H) and two identical light chains (L), each with variable (V) and constant (C) domains. The antigen binding sites are situated in the arms of the molecule and are comprised of the CDR loops of the variable V_H and V_L domains. Framework regions are located between these CDRs. The hinge and the interchain disulfide bonds connecting the heavy chains or the heavy and light chain at C_L and C_H1, respectively, are indicated in green. Green lines also represent intrachain disulfide bonds. The conserved glycosylation at Asn297 of the C_H2 domain is shown in red. Papain digest generates the two antigen binding Fab and one Fc fragment indicated in this scheme adapted from Beck and colleagues.[7] Abbreviations: C_L: constant domain of the light chain; C_H1-C_H3: constant domains of the heavy chain; V_L: variable domain of the light chain; V_H: variable domain of the heavy chain.

As depicted in Fig. 2.1, each of the IgG chains consist of multiple domains (two or four for the light and heavy chain, respectively), all possessing the so-called "immunoglobulin fold". This characteristic structural unit has a molecular weight of ~12.5 kDa (corresponding to about 110 amino acids) and consists of two antiparallel β-sheets forming a β-sandwich (Fig. 2.2a).[8; 9] In constant domains, this β-sandwich is composed of seven β-strands while two additional strands are present in the variable domains. A hallmark of this immunoglobulin fold is the intradomain disulfide bond that covalently links the two β-sheets and is deeply buried in the interior of the IgG domain (Fig. 2.2b).

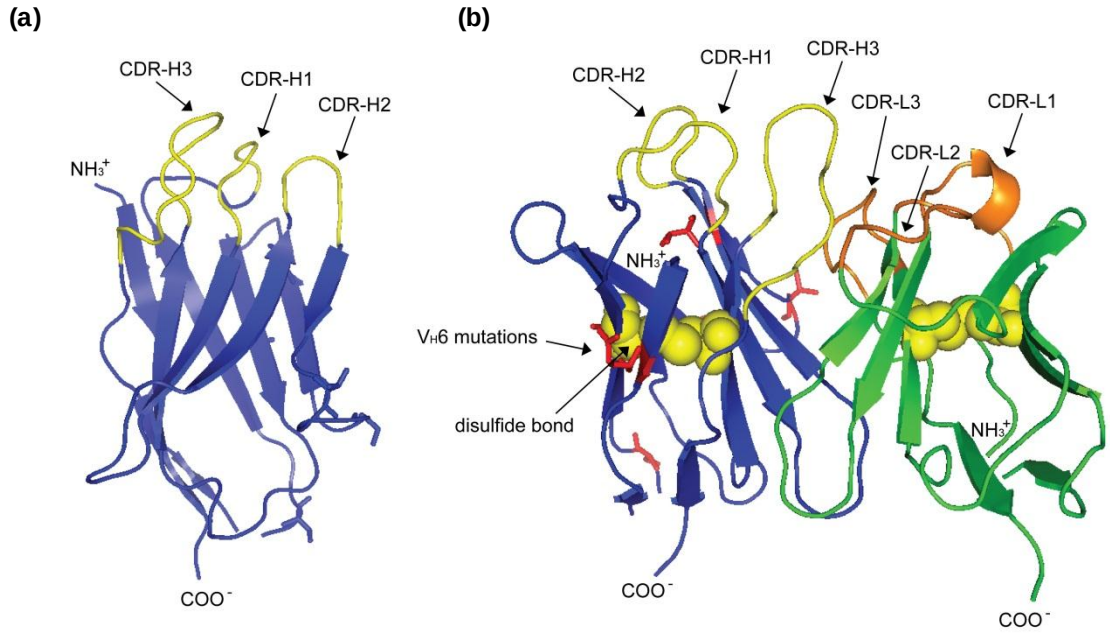


Fig. 2.2: Topological structure of the immunoglobulin fold and the 6B3 Fv fragment. **(a)** The model structure of V_H6 was used to illustrate the nine-stranded antiparallel β-barrel of a variable domain. The three complementarity determining regions (CDR-H1 to 3, indicated in yellow) protrude as loops from the variable domain core structure and form half of the antigen binding site. **(b)** Model of the 6B3 Fv fragment (PDB entry VL3_VH6_FV). The V_H domain is shown in blue with yellow CDR loops while the V_L domain is depicted in green with orange CDRs. Yellow spheres represent both intrachain disulfide bonds. Residues modified within this study (see below) are highlighted as red sticks.

The immune system of vertebrates relies on antibodies to protect the organism against infections by various pathogens. Based on the different structures of the heavy chains, human immunoglobulins can be grouped into the five main classes, also termed isotypes, IgA, IgD, IgE, IgG and IgM with their corresponding heavy chains α, δ, ε, γ and μ. These classes differ in their biological properties, their functional locations and the ability to deal with different antigens. In the early stages of B-cell mediated immunity, IgM molecules are produced to eliminate pathogens. Secreted members of this isotype are polymeric proteins with a starlike structure composed of five single Ig subunits, thereby revealing a rather high avidity to their targets. IgA (with human subclasses IgA1 and IgA2) is present in mucosal areas preventing their colonization by pathogens and has both a monomeric variant present in the serum and a dimeric secreted form. All other isotypes - IgD, IgE and IgG (with its four human subclasses IgG1-IgG4, see below) - are only to be found in the serum as monomers. Of these immunoglobulins, IgG provides the majority of antibody-based immunity against invading pathogens and is the only type of antibody that can cross the placenta during pregnancy due to its binding to placental trophoblasts. Importantly, while antibodies of all other isotypes have a relatively short half-life of only a couple of days, IgGs have a much longer half-life of up to three weeks, making them a very important component of the human immune system.

The human IgG family itself consists of four distinct subclasses (IgG1, IgG2, IgG3 and IgG4, respectively) which are known to possess differences in certain antigenic, structural and biologic properties.[10] Having a sequence homology of more than 95%, the subclasses show their most conspicuous differences in the amino acid composition and structure of their hinge regions with their interchain disulfide bonds.[11] Thus, the length and the flexibility of the hinge regions vary considerably

among the different subclasses. While e.g. the hinge of IgG1 can move rather freely, its flexibility is restricted for IgG2 and IgG4 due to its shorter length.[12] IgG3, on the contrary, has a peculiar extended hinge region and exhibits eleven interchain disulfide bonds – substantially more than the two or four disulfide bonds found in IgG1 and IgG4 or IgG 2, respectively. Irrespective of these differences, the Fc regions of each IgG subclass still are structurally very similar with distinct affinities for the Fc receptors. However, IgG1 and IgG3 show a more effective binding than IgG2 and IgG4.[10]

Diversity of antibodies

To fulfill its task, the human immune system requires a large repertoire of immunoglobulin molecules, differing in the sequence and the structure of their variable regions. This diversity is achieved through a combination of various rearrangements and combinatorial assemblies of different genetic segments.[13] The variable domains of the heavy chain are generated by the combination of variable (V), diversity (D) and joining (J) gene segments while those of the light chain only consist of V- and J-regions. The variable V_H domain is an assembly of a segment of the approximately 50 functional V-segments, 30 D-segments and 6 J-segments each, all located on chromosome 14.[14] For the κ-light chains, around 40 functional V-segments and 5 J-segments are located on chromosome 2[15] whereas 31 V-segments as well as 6 J-segments on chromosome 22 are used to create the variable domain of the λ counterpart.[16] The arrangement of these gene segments, termed V(D)J recombination, takes place during the development of B-lymphocytes.[17-19] The diversity arising from these combination events is even further increased through nucleotide deletions or insertions at the junction sites of the V, D and J fragments[17] as well as through the random combination of a certain heavy with any other light chain.

The last diversification mechanism occurs after B-cell maturation is completed. When first contacts with the antigen have been made, stimulated B-lymphocytes start a process termed somatic hypermutation[20] in which several point mutations are introduced into the variable regions of the rearranged genes. This last process, also called affinity maturation, increases the potential and effective diversity by many orders of magnitude. All of the mentioned mechanisms finally lead to a theoretical diversity of up to 10^{12} different immunoglobulins, which are subsequently expressed on the surface of B-cells and secreted as soluble antibodies into the serum.[21]

Selection of recombinant antibodies and different immunoglobulin formats

Next to the isolation of polyclonal antibodies from immunized animals and the hybridoma technology developed by Köhler and Milstein,[22] the recombinant construction of antibodies is the most popular method for obtaining immunoglobulins nowadays. The knowledge of the IgG structure in molecular details and the development of advanced methods in gene technology over the last decades significantly contributed to the success of this technology. Today, numerous libraries have been described[23] and antibodies can be obtained from naive repertoires[24; 25] or libraries of fully synthetic genes[26] - allowing the selection of recombinant antibodies of completely human nature with regard to their sequence and characteristics.

A requirement for the selection of antibodies from man-made libraries was the development and establishment of the phage display technology.[27; 28] Thus, it became possible to select specific binders from large recombinant *in vitro* libraries. Nowadays, there are also alternative display technologies that couple the genotype to the phenotype of the selected binders such as yeast display,[29] bacterial display[30] and further *in vitro* display technologies including ribosome display.[31] In most of these techniques, antibodies are not selected as full-length IgGs but rather are used in a variety of smaller formats. In the past, several antibody fragments have been created of which some are applicable for these selection methods. Generally, the Fv-fragment[32] (V_L/V_H; shown on top of Fig. 2.3) is the minimal setup which still contains the complete antigen binding site. However, due to the non-covalent connection of both variable domains, this fragment is generally rather instable and not well suited for selection purposes. This problem was solved by the development of the scFv (single chain Fv; depicted on the left side of Fig. 2.3) fragments,[33; 34] single polypeptides containing an additional 15 to 20 amino acids long flexible peptide linker connecting the V_H with the V_L domain (or vice versa). As an alternative, Fv fragments could be stabilized by the introduction of a disulfide bond between the C-terminal regions of the variable domains – leading to so-called dsFvs fragments which, however, shows lower expression levels than scFvs.[35] Next to scFvs, also Fab-fragments ("fragment antigen binding"; illustrated on the right side of Fig. 2.3) are frequently used in selection setups. These fragments are heterodimers and consist of the complete light chain (V_L/C_L) and the Fd-fragment (V_H/C_H1) of the heavy chain.

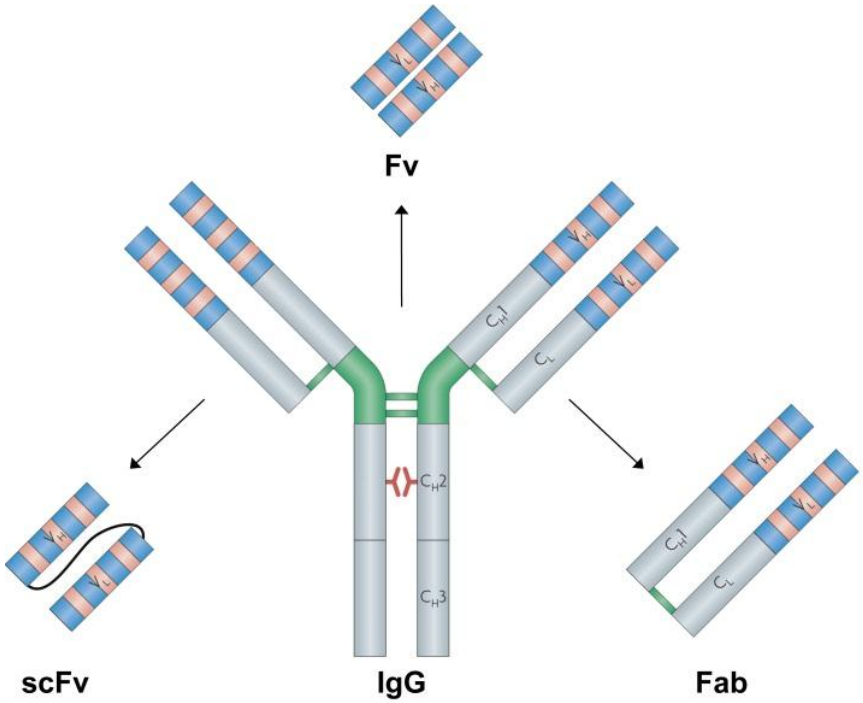


Fig. 2.3: Model structures of various antibody fragments. Derivatives of the full-length IgG molecule, indicated in the centre. The scFv fragment is depicted on the left side, consisting of V_L and V_H connected by a flexible peptide linker. Shown is only the V_H-V_L orientation although the other orientation also exists. The Fab-fragment, displayed on the right side, is composed of the whole light chain and the Fd fragment of the heavy chain (V_H and C_H1 domain) connected by a disulfide bond between C_L and C_H1. Above the IgG, the unlinked Fv fragment is shown.

Due to their small size of ~30 kDa and the relative ease with which these single polypeptide chains can be expressed in bacteria, scFv fragments are a very popular, if not the most popular antibody format.[36] Compared to full-length IgGs, they possess some clinical advantages such as better tumor penetration[37] and faster clearance from the organism.[38] However, neither scFv nor Fab fragments contain any effector function part which could activate compounds of the immune system, as described above. Although there are some applications which do not depend on or even want to prevent any immune response,[39; 40] most marketed antibodies still comprise full-length human IgG molecules, mainly since long half-lives have been considered desirable for most applications.

Production of full-length IgGs and antibody fragments

During the mid-1980's, developments in molecular biology and recombinant DNA techniques allowed the first expression of functional full-length antibodies in yeast[41] and cells of higher eukaryotes.[42; 43] Since the yields were rather low and the cultivation of these cells quite elaborate, antibody expression systems were also developed for the bacterium *Escherichia coli*.[44; 45] However, it was not until the genes for the heavy and light chain, respectively, were fused to N-terminal signal sequences that functional antibody fragments were successfully expressed in *E. coli*.[32; 46] These signal sequences allowed the secretion of the antibody fragments into the periplasm, whose oxidative environment and various chaperones allowed the proper folding and formation of disulfide bonds which are important for the structure of the immunoglobulin domains.[47] Ever since, bacterial hosts have been commonly used for the production of antibody fragments, mainly due to the ease of genetic manipulation of this well established and relatively inexpensive expression system. However, while the expression of small antibody fragments (like scFv or Fab fragments) in *E. coli* nowadays can be performed at large scale in fermenter cultures,[48; 49] often only a minor fraction of product can be recovered in functional form due to limitations of the antibody assembly in the periplasmic space. Even though meanwhile the expression of full-length IgG has been achieved in *E. coli*,[50-52] their levels are still very low and incompletely assembled molecules also occur. Therefore, therapeutic antibodies generally are produced in mammalian cells, since in addition only this expression system can carry out the proper posttranslational modifications necessary for the functional activities of IgGs.

To date, different expression systems have been developed for the recombinant production of various antibody formats. In addition to other prokaryotic systems such as *Bacillus subtillis*,[53] *Streptomyces lividans*[54] or *Proteus mirabilis*,[55] several eukaryotic systems were successfully employed, such as yeasts and filamentous fungi,[56; 57] insect cells,[58] algae,[59] plants[60] and transgenetic animals.[61] However, currently all approved therapeutic antibodies of the IgG format are produced in mammalian cells,[62] since only this system introduces the complex glycosylation which is necessary for the interaction of the antibodies with Fcγ receptors.[63] In addition, mammalian cells can be cultured in large-scale fermenters, offer high expression yields and can be optimized due to recent improvements in cell engineering.[64; 65] Furthermore, nowadays the establishment of stable cell lines, expressing the antibodies permanently at high levels from constitutively active promoters, can be easily achieved.[64; 66] Of additional importance are the sophisticated folding machinery and quality control mechanisms, located in the endoplasmatic reticulum of mammalian cells.[67; 68]

In some cases, however, the FcγR interaction is neither needed nor wanted, thus widening the range of potential production hosts to e.g. *Pichia pastoris*. In the past, this methylotrophic yeast has been widely used for the expression of both secreted and intracellular eukaryotic proteins.[69] *P. pastoris* was initially developed as a biological tool in the 1970's to exploit methanol, the byproduct of the oil industry, to generate yeast biomass or single cell protein which could be used as livestock feed. Being a microbial eukaryote, *Pichia* combines several advantages of eukaryotic expression systems, like its capacity of performing posttranslational modifications, with the easy handling and simplicity of techniques needed for the genetic manipulation of microbes like *E. coli*. In addition, *P. pastoris* has a shorter generation cycle and is cheaper to use than mammalian expression systems. While it is able to carry out efficient disulfide bond formation and isomerization using the eukaryotic secretion quality control machinery,[69-71] *Pichia* expression generates mannose-rich glycosylation of IgG.[72] The attached glycan moieties differ markedly from the product of mammalian cells, resulting in antibodies that lack FcγR binding capacity. However, recently, *P. pastoris* has been engineered to introduce complex, human-like glycosylation.[73-76]

Due to its respiratory growth, *P. pastoris* can grow to very high cell densities,[77] and therefore express high levels of recombinant protein[78] through efficient and tightly regulated promoters, accounting for up to 30% of the total cell protein.[79] Moreover, comparatively few endogenous proteins are secreted by this yeast system, which offers further advantages for efficient subsequent protein purification and downstream processing.[80] Since Ridder and colleagues reported the first successful production of functional antibody fragments in *Pichia*,[81] more than 50 reports have been published describing antibody expression in this system.[82] While most of these molecules were scFv fragments and fusions thereof, several Fab fragments have been produced in *Pichia* as well.[83-85] Interestingly, until today *P. pastoris* has only been used as an expression host for a handful of full-length IgGs.[75; 76; 86]

Antibody Engineering

The availability of a number of crystal structures and their modular composition make immunoglobulins quite suitable for protein engineering attempts. In recent years, two principal methods have proven to be successful for improving antibody sequences, being classifiable as either "rational" approaches or "directed evolution" ones.

Rational or structure-based engineering relies on the detailed analysis of three-dimensional (3D) structures of the protein of interest (POI) and usually results in the exchange of certain amino acids by site-directed mutagenesis.[87] To optimize key residue interactions, at first those suboptimal amino acids in the original sequences leading to decreased stability have to be determined, followed by the careful decision with which amino acid they should be replaced. Therefore, the rational approach partly is based on alignments of particular antibody sequences to well-expressing stable and non-aggregating fragments,[88-92] an analysis of exposed hydrophobic residues[93] or the grafting of CDRs onto a stable and well folding framework.[94-96] Thereby, several aspects such as the packing of the hydrophobic core, hydrogen bonding or electrostatic interactions among many others must be taken into account.[97] To avoid unfavorable strain in the resulting mutants, high-resolution structures are a fundamental prerequisite to allow the estimation of possible conformational as well as energetic and steric

influences upon the replacements. With the present efforts in the field of structural genomics, these rational approaches are likely to become more and more important in the future.

On the other hand, different evolutionary strategies have been designed, all of them mimicking principles of Darwinian evolution.[97] In this approach, the POI is subjected to an evolutionary pressure which often rewards stability and expression.[97-99] After creating a randomized genetic library of mutants derived from the wild-type sequence of the POI, sequence diversity can be increased by several methods. Using for example either error-prone PCR,[100] bacterial mutator strains[101] or DNA shuffling,[102] random mutations are introduced into the sequence and the resulting proteins subsequently screened and selected under defined conditions for the desired features. For selection, the phenotype must be linked to its genotype in order to use the positive-selected members for the next round of evolution. As mentioned above, phage and ribosome display have proven to be powerful evolutionary technologies which can be used for the optimization of protein structure and interactions.[28; 31]

Combining these two approaches, several different methods for antibody engineering have proven to be successful[97] and have gained more and more importance over the last few years. Recent advances in the engineering of protein stability have shown that even minor changes in a given sequence can have profound effects on the biochemical and biophysical properties of the respective protein. Thus, for many proteins, some specific mutations may be sufficient enough to considerably increase these characteristic features. Hence, the challenge of protein engineering is to identify these positions in a given sequence among the vast number of possible changes and to advantageously exchange them.

Context and aim of this study

Due to their high degree of specificity and broad target range, recombinant antibodies are used in an ever increasing number of applications. These make great demands on the molecules concerning their obtainable production yield, a low aggregation-tendency and their stability against high denaturant concentrations, proteases and elevated temperatures. However, only a subset of antibodies actually has the biophysical properties required for these applications. The resulting inability to reach the disease-relevant target in sufficient concentrations is one of the biggest problems limiting the clinical usage of recombinant antibodies.[103] Therefore, engineering approaches nowadays are particularly aiming at the improvement of the expression levels and overall stabilities of IgG molecules. This is of great importance, as the high costs of production are a further major drawback of these molecules. More stable antibodies would result in longer half-lives *in vivo* and thereby decrease both the dosage as well as the frequency at which antibody drugs have to be administered. In addition, the best binders are useless if they cannot be produced in sufficient amounts or if they are not stable enough to fulfill their tasks.

Previously, our laboratory as well as others analyzed sequence modifications in the variable domain for their impact on aggregation propensity and conformational stability.[90; 93; 104-106] By analyzing the human antibody repertoire in terms of structure, amino acid sequence diversity and germline usage, our laboratory showed that seven V_H (V_H1a, V_H1b, V_H2, V_H3, V_H4, V_H5, V_H6) and seven V_L (V_κ1, V_κ2,

$V_\kappa 3$ $V_\kappa 4$ and $V_\lambda 1$, $V_\lambda 2$, $V_\lambda 3$) germline families cover more than 95% of the human antibody diversity.[26] Based on these data, consensus framework sequences could be derived for each of these families, eventually resulting in the design of the fully synthetic Human Combinatorial Antibody Library (HuCAL).[26] These subfamily-specific consensus sequences cover 98% of all V_H, 99% of all V_κ and 93% of all V_λ sequences. The biophysical properties of these V_H and V_L consensus domains were subsequently evaluated alone and in different V_H/V_L combinations in the scFv context for their expression behavior and denaturant-induced unfolding equilibria under non-reducing conditions.[107] This study revealed that antibody fragments containing V_H domains of the even-numbered human germline families 2, 4 and 6 had comparatively lower expression yields as well as lower midpoints and low cooperativity in equilibrium unfolding compared to those of the odd-numbered families 1a, 1b, 3 and 5.

In subsequent studies,[91] members of our laboratory performed sequence comparisons of the seven V_H consensus domains (see Fig. 2.4), detecting residues potentially responsible for the unsatisfactory features of the families with less favorable biophysical properties (V_H2, V_H4, V_H6). These sequence analyses revealed differences in the packing of the hydrophobic core, the completeness of charge clusters, the occurrence of unsatisfied hydrogen bonds and residues with low β-sheet propensities, positive Φ angles as well as exposed hydrophobic side chains. Interestingly, most of these amino acid residues were present in all members of the even-numbered families (V_H1a, V_H1b, V_H3, V_H5), but did not occur in the odd-numbered ones. Based on this sequence comparison in conjunction with analyses of structural models, six distinct residues of the framework regions were chosen which were characteristic for their respective family and which were hypothesized to be acting independently and to be individually exchangeable. Among the performed mutations was the replacement of non-glycine loop residues with positive φ angles to glycines (positions 16 and 76, numbering scheme according to Honegger and Plückthun[5]), as well as the exchange of non-optimal β-strand residues with amino acids possessing higher β-sheet propensity (residues 5 and 90). In addition, at position 72, a solvent-exposed hydrophobic residue was mutated into a hydrophilic one and residue 58 having an unsatisfied H-bond was replaced. All the selected mutations (Q5V, S16G, T58I, V72D, S76G and S90Y, highlighted in Fig. 2.4 by red bold letters in the columns with the blue background) were located only in the V_H framework and not in the CDR regions, since these mutations were intended to be as generally applicable as possible and unlikely to affect antigen binding. The exact location of these mutations within the Fv fragment can be seen in Fig. 2.3b with the mutated residues highlighted in red.

Indeed, the introduction of the six mutated residues turned out to dramatically increase both the expression yield as well as the thermodynamic stability of engineered scFv fragments produced in the prokaryotic system of *E. coli* without affecting their binding properties. These data were recorded for two model antibody scFv named 2C2 and 6B3 that had originally been selected from the HuCAL library by panning against the peptide M18 coupled to transferrin or myoglobin from horse skeletal muscle, respectively. These molecules were chosen as both belonged to the human V_H6 framework family whose members are known to have a rather aggregation prone behavior and the lowest thermodynamic stability of all human V_H domains.[107] Originating from the same framework family, their V_H6 sequences were identical except for different CDR-H3 loops, causing the 2C2 and 6B3 constructs to recognize different antigens. In addition, the two antibodies differed in the nature of their respective

light chain, the 2C2 having a $V_κ3$ and 6B3 containing a $V_λ3$ light chain. Importantly, the improvements found for the engineered scFv fragments were independent of both, the V_L domain and the sequence and length of the CDR-H3, as both constructs gave similar results concerning increased scFv yields and stabilities.

In this project and in most studies of other laboratories, the vast majority of antibody engineering attempts are performed for antibody formats smaller than the full-length IgG. This was done in order to obtain interpretable results, as the occurrence of multiple domains in the IgG make the analysis of unfolding curves extremely difficult. A further complication arises from the interaction of the domains.[36] In particular, scFv fragments expressed in *E. coli* have been the focus of most studies concerning biochemical and biophysical properties so far. However, so far very little is known about the extent with which results derived from studies using a certain antibody format (like scFv) can be applied to other antibody formats (like full-length IgGs). Keeping in mind that most approved therapeutic antibodies are in the IgG format and produced in mammalian cells, the engineering of full-length antibodies will become more and more important in the years to come. Unfortunately, almost nothing is known about the biochemical and biophysical effects of mutations in the context of the IgG molecule made in mammalian cells. Therefore, the main focus of this thesis was to analyze whether the previously found beneficial mutations in scFvs still stayed true in the context of "less-truncated" antibody-formats (Fab fragment and full-length IgGs) and to investigate the influence of the expression system.

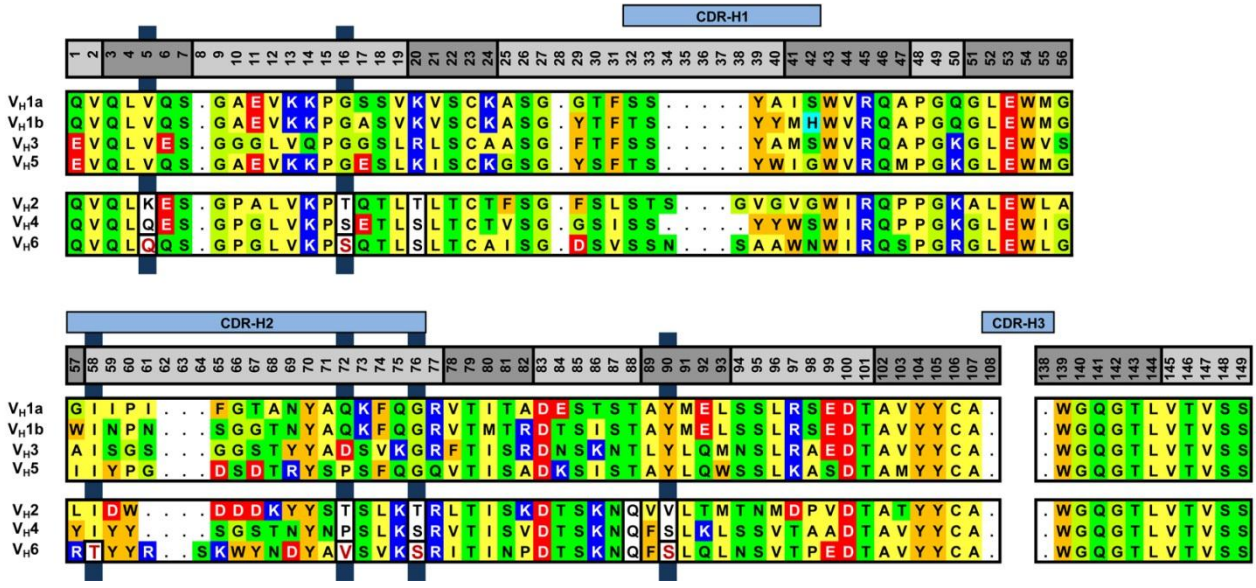


Fig. 2.4: Aligned HuCAL V_H sequences. The amino acids are color-coded according to residue type: aromatic residues (Tyr, Phe, Trp) are orange, hydrophobic residues (Leu, Ile, Val, Met, Cys, Pro, Ala) yellow, uncharged hydrophilic residues (Ser, Thr, Gln, Asn, Gly) green, acidic residues (Asp, Glu) red and basic residues (Arg, Lys, His) blue. Residues that show correlated sequence differences among the groups of V_H domains with favorable properties (V_H1a, V_H1b, V_H3, and V_H5) and V_H domains with less favorable properties (V_H2, V_H4, and V_H6) are indicated by red bold letters in the columns with the dark blue background. The numbering scheme is according to Honegger and Plückthun[5] and the figure adapted from Ewert and collegues.[91]

A set of different expression hosts (*E. coli*, *P. pastoris* and different mammalian cells) was established as production systems. Thus, it should be possible to determine the influences of the respective protein folding and quality control machineries, while the use of different antibody formats should give insight into the influence of intrinsic domain interactions. To test the transferability of the previous results and their to different antibody formats, both the wild-type (WT) and the best engineered variant (called M throughout this work) were converted to the Fab or full-length IgG format and investigated for their expression characteristics and biophysical properties. In addition, the established expression systems were used to distinguish how various cellular hosts deal with aggregation-prone proteins in general and to compare the biophysical properties of the corresponding antibodies. As one major challenge encountered in the development of antibody-based therapeutics is their aggregation susceptibility under formulation conditions[108; 109] and in vivo, special attention was paid to this feature. Since irreversible aggregation generally leads to reduced bioactivity[110] and presents an increased risk of immunogenicity,[111-116] engineering IgG to reduce or eliminate its self-association is of great importance and might further extend the potential of therapeutic antibodies.

References

1. Maynard, J. & Georgiou, G. (2000). Antibody engineering. *Annu Rev Biomed Eng* **2**, 339-376.
2. Farid, S. S. (2006). Established bioprocesses for producing antibodies as a basis for future planning. *Adv Biochem Eng Biotechnol* **101**, 1-42.
3. Carter, P. (2001). Improving the efficacy of antibody-based cancer therapies. *Nat Rev Cancer* **1**, 118-129.
4. Kabat, E. A., Wu, T. T. & Bilofsky, H. (1977). Unusual distributions of amino acids in complementarity-determining (hypervariable) segments of heavy and light chains of immunoglobulins and their possible roles in specificity of antibody-combining sites. *J Biol Chem* **252**, 6609-6616.
5. Honegger, A. & Plückthun, A. (2001). Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling and analysis tool. *J Mol Biol* **309**, 657-670.
6. Jefferis, R., Lund, J. & Pound, J. D. (1998). IgG-Fc-mediated effector functions: molecular definition of interaction sites for effector ligands and the role of glycosylation. *Immunol Rev* **163**, 59-76.
7. Beck, A., Wurch, T., Bailly, C. & Corvaia, N. (2010). Strategies and challenges for the next generation of therapeutic antibodies. *Nat Rev Immunol* **10**, 345-352.
8. Poljak, R. J., Amzel, L. M., Avey, H. P., Chen, B. L., Phizackerley, R. P. & Saul, F. (1973). Three-dimensional structure of the Fab' fragment of a human immunoglobulin at 2,8-Å resolution. *Proc Natl Acad Sci U S A* **70**, 3305-3310.
9. Lesk, A. M. & Chothia, C. (1982). Evolution of proteins formed by beta-sheets. II. The core of the immunoglobulin domains. *J Mol Biol* **160**, 325-342.
10. Jefferis, R., Pound, J., Lund, J. & Goodall, M. (1994). Effector mechanisms activated by human IgG subclass antibodies: clinical and molecular aspects. *Ann. Biol. Clin. (Paris)* **52**, 57-65.
11. Pan, Q. & Hammarstrom, L. (2000). Molecular basis of IgG subclass deficiency. *Immunol. Rev.* **178**, 99-110.
12. Burton, D. R., Gregory, L. & Jefferis, R. (1986). Aspects of the molecular structure of IgG subclasses. *Monogr. Allergy* **19**, 7-35.
13. Rolink, A. & Melchers, F. (1993). Generation and regeneration of cells of the B-lymphocyte lineage. *Curr Opin Immunol* **5**, 207-217.
14. Cook, G. P. & Tomlinson, I. M. (1995). The human immunoglobulin VH repertoire. *Immunol Today* **16**, 237-242.
15. Tomlinson, I. M., Cox, J. P., Gherardi, E., Lesk, A. M. & Chothia, C. (1995). The structural repertoire of the human V kappa domain. *EMBO J* **14**, 4628-4638.

16. Frippiat, J. P., Williams, S. C., Tomlinson, I. M., Cook, G. P., Cherif, D., Le Paslier, D., Collins, J. E., Dunham, I., Winter, G. & Lefranc, M. P. (1995). Organization of the human immunoglobulin lambda light-chain locus on chromosome 22q11.2. *Hum Mol Genet* **4**, 983-991.
17. Gauss, G. H. & Lieber, M. R. (1996). Mechanistic constraints on diversity in human V(D)J recombination. *Mol Cell Biol* **16**, 258-269.
18. Grawunder, U., West, R. B. & Lieber, M. R. (1998). Antigen receptor gene rearrangement. *Curr Opin Immunol* **10**, 172-180.
19. Hozumi, N. & Tonegawa, S. (1976). Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc Natl Acad Sci U S A* **73**, 3628-3632.
20. Wagner, S. D. & Neuberger, M. S. (1996). Somatic hypermutation of immunoglobulin genes. *Annu Rev Immunol* **14**, 441-457.
21. Berek, C. & Ziegner, M. (1993). The maturation of the immune response. *Immunol Today* **14**, 400-404.
22. Köhler, G. & Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**, 495-497.
23. Mondon, P., Dubreuil, O., Bouayadi, K. & Kharrat, H. (2008). Human antibody libraries: a race to engineer and explore a larger diversity. *Front Biosci* **13**, 1117-1129.
24. Winter, G., Griffiths, A. D., Hawkins, R. E. & Hoogenboom, H. R. (1994). Making antibodies by phage display technology. *Annu Rev Immunol* **12**, 433-455.
25. Vaughan, T. J., Williams, A. J., Pritchard, K., Osbourn, J. K., Pope, A. R., Earnshaw, J. C., McCafferty, J., Hodits, R. A., Wilton, J. & Johnson, K. S. (1996). Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nat Biotechnol* **14**, 309-314.
26. Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wölle, J., Plückthun, A. & Virnekäs, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J Mol Biol* **296**, 57-86.
27. Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315-1317.
28. Smith, G. P. & Petrenko, V. A. (1997). Phage Display. *Chem Rev* **97**, 391-410.
29. Boder, E. T. & Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol* **15**, 553-557.
30. Francisco, J. A., Campbell, R., Iverson, B. L. & Georgiou, G. (1993). Production and fluorescence-activated cell sorting of Escherichia coli expressing a functional antibody fragment on the external surface. *Proc Natl Acad Sci U S A* **90**, 10444-10448.
31. Hanes, J. & Plückthun, A. (1997). *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc. Natl. Acad. Sci. U S A* **94**, 4937-4942.
32. Skerra, A. & Plückthun, A. (1988). Assembly of a functional immunoglobulin Fv fragment in Escherichia coli. *Science* **240**, 1038-1041.
33. Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R. & et al. (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. *Proc Natl Acad Sci U S A* **85**, 5879-5883.
34. Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S. M., Lee, T., Pope, S. H., Riordan, G. S. & Whitlow, M. (1988). Single-chain antigen-binding proteins. *Science* **242**, 423-426.
35. Glockshuber, R., Malia, M., Pfitzinger, I. & Plückthun, A. (1990). A comparison of strategies to stabilize immunoglobulin Fv-fragments. *Biochemistry* **29**, 1362-1367.
36. Röthlisberger, D., Honegger, A. & Plückthun, A. (2005). Domain interactions in the Fab fragment: a comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability. *J Mol Biol* **347**, 773-789.
37. Yokota, T., Milenic, D. E., Whitlow, M. & Schlom, J. (1992). Rapid tumor penetration of a single-chain Fv and comparison with other immunoglobulin forms. *Cancer Res* **52**, 3402-3408.
38. Kang, N., Hamilton, S., Odili, J., Wilson, G. & Kupsch, J. (2000). In vivo targeting of malignant melanoma by 125Iodine- and 99mTechnetium-labeled single-chain Fv fragments against high molecular weight melanoma-associated antigen. *Clin Cancer Res* **6**, 4921-4931.
39. Bolt, S., Routledge, E., Lloyd, I., Chatenoud, L., Pope, H., Gorman, S. D., Clark, M. & Waldmann, H. (1993). The generation of a humanized, non-mitogenic CD3 monoclonal antibody which retains in vitro immunosuppressive properties. *Eur J Immunol* **23**, 403-411.

40. Reddy, M. P., Kinney, C. A., Chaikin, M. A., Payne, A., Fishman-Lobell, J., Tsui, P., Dal Monte, P. R., Doyle, M. L., Brigham-Burke, M. R., Anderson, D., Reff, M., Newman, R., Hanna, N., Sweet, R. W. & Truneh, A. (2000). Elimination of Fc receptor-dependent effector functions of a modified IgG4 monoclonal antibody to human CD4. *J Immunol* **164**, 1925-1933.
41. Wood, C. R., Boss, M. A., Kenten, J. H., Calvert, J. E., Roberts, N. A. & Emtage, J. S. (1985). The synthesis and in vivo assembly of functional antibodies in yeast. *Nature* **314**, 446-449.
42. Morrison, S. L., Johnson, M. J., Herzenberg, L. A. & Oi, V. T. (1984). Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains. *Proc Natl Acad Sci U S A* **81**, 6851-6855.
43. Roberts, S. & Rees, A. R. (1986). The cloning and expression of an anti-peptide antibody: a system for rapid analysis of the binding properties of engineered antibodies. *Protein Eng* **1**, 59-65.
44. Cabilly, S., Riggs, A. D., Pande, H., Shively, J. E., Holmes, W. E., Rey, M., Perry, L. J., Wetzel, R. & Heyneker, H. L. (1984). Generation of antibody activity from immunoglobulin polypeptide chains produced in Escherichia coli. *Proc Natl Acad Sci U S A* **81**, 3273-3277.
45. Boss, M. A., Kenten, J. H., Wood, C. R. & Emtage, J. S. (1984). Assembly of functional antibodies from immunoglobulin heavy and light chains synthesised in E. coli. *Nucleic Acids Res* **12**, 3791-3806.
46. Better, M., Chang, C. P., Robinson, R. R. & Horwitz, A. H. (1988). Escherichia coli secretion of an active chimeric antibody fragment. *Science* **240**, 1041-1043.
47. Glockshuber, R., Schmidt, T. & Plückthun, A. (1992). The disulfide bonds in antibody variable domains: effects on stability, folding in vitro, and functional expression in Escherichia coli. *Biochemistry* **31**, 1270-1279.
48. Harrison, J. S. & Keshavarz-Moore, E. (1996). Production of antibody fragments in Escherichia coli. *Ann N Y Acad Sci* **782**, 143-158.
49. Humphreys, D. P. (2003). Production of antibodies and antibody fragments in Escherichia coli and a comparison of their functions, uses and modification. *Curr Opin Drug Discov Devel* **6**, 188-196.
50. Simmons, L. C., Reilly, D., Klimowski, L., Raju, T. S., Meng, G., Sims, P., Hong, K., Shields, R. L., Damico, L. A., Rancatore, P. & Yansura, D. G. (2002). Expression of full-length immunoglobulins in Escherichia coli: rapid and efficient production of aglycosylated antibodies. *J Immunol Methods* **263**, 133-147.
51. Makino, T., Skretas, G., Kang, T. H. & Georgiou, G. (2011). Comprehensive engineering of Escherichia coli for enhanced expression of IgG antibodies. *Metab Eng* **13**, 241-251.
52. Mazor, Y., Van Blarcom, T., Mabry, R., Iverson, B. L. & Georgiou, G. (2007). Isolation of engineered, full-length antibodies from libraries expressed in Escherichia coli. *Nat Biotechnol* **25**, 563-565.
53. Wu, X. C., Ng, S. C., Near, R. I. & Wong, S. L. (1993). Efficient production of a functional single-chain antidigoxin antibody via an engineered Bacillus subtilis expression-secretion system. *Biotechnology (N Y)* **11**, 71-76.
54. Ueda, Y., Tsumoto, K., Watanabe, K. & Kumagai, I. (1993). Synthesis and expression of a DNA encoding the Fv domain of an anti-lysozyme monoclonal antibody, HyHEL10, in Streptomyces lividans. *Gene* **129**, 129-134.
55. Rippmann, J. F., Klein, M., Hoischen, C., Brocks, B., Rettig, W. J., Gumpert, J., Pfizenmaier, K., Mattes, R. & Moosmayer, D. (1998). Procaryotic expression of single-chain variable-fragment (scFv) antibodies: secretion in L-form cells of Proteus mirabilis leads to active product and overcomes the limitations of periplasmic expression in Escherichia coli. *Appl Environ Microbiol* **64**, 4862-4869.
56. Joosten, V., Lokman, C., Van Den Hondel, C. A. & Punt, P. J. (2003). The production of antibody fragments and antibody fusion proteins by yeasts and filamentous fungi. *Microb Cell Fact* **2**, 1.
57. Gasser, B. & Mattanovich, D. (2007). Antibody production with yeasts and filamentous fungi: on the road to large scale? *Biotechnol Lett* **29**, 201-212.
58. Reavy, B., Ziegler, A., Diplexcito, J., Macintosh, S. M., Torrance, L. & Mayo, M. (2000). Expression of functional recombinant antibody molecules in insect cell expression systems. *Protein Expr Purif* **18**, 221-228.
59. Franklin, S. E. & Mayfield, S. P. (2005). Recent developments in the production of human therapeutic proteins in eukaryotic algae. *Expert Opin Biol Ther* **5**, 225-235.
60. Nölke, G., Fischer, R. & Schillberg, S. (2003). Production of therapeutic antibodies in plants. *Expert Opin Biol Ther* **3**, 1153-1162.
61. Young, M. W., Meade, H., Curling, J. M., Ziomek, C. A. & Harvey, M. (1998). Production of recombinant antibodies in the milk of transgenic animals. *Res Immunol* **149**, 609-610.

62. Andersen, D. C. & Krummen, L. (2002). Recombinant protein expression for therapeutic applications. *Curr Opin Biotechnol* **13**, 117-123.
63. Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M. & Dwek, R. A. (2007). The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu. Rev. Immunol.* **25**, 21-50.
64. Barnes, L. M. & Dickson, A. J. (2006). Mammalian cell factories for efficient and stable protein expression. *Curr Opin Biotechnol* **17**, 381-386.
65. Omasa, T., Onitsuka, M. & Kim, W. D. (2010). Cell engineering and cultivation of chinese hamster ovary (CHO) cells. *Curr Pharm Biotechnol* **11**, 233-240.
66. Zhou, H., Liu, Z. G., Sun, Z. W., Huang, Y. & Yu, W. Y. (2010). Generation of stable cell lines by site-specific integration of transgenes into engineered Chinese hamster ovary strains using an FLP-FRT system. *J Biotechnol* **147**, 122-129.
67. Kleizen, B. & Braakman, I. (2004). Protein folding and quality control in the endoplasmic reticulum. *Curr Opin Cell Biol* **16**, 343-349.
68. Brodsky, J. L. & Skach, W. R. (2011). Protein folding and quality control in the endoplasmic reticulum: Recent lessons from yeast and mammalian cell systems. *Curr Opin Cell Biol* **23**, 464-475.
69. Cregg, J. M., Tolstorukov, I., Kusari, A., Sunga, J., Madden, K. & Chappell, T. (2009). Expression in the yeast Pichia pastoris. *Methods Enzymol.* **463**, 169-189.
70. Cereghino, G. P., Cereghino, J. L., Ilgen, C. & Cregg, J. M. (2002). Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris. *Curr. Opin. Biotechnol.* **13**, 329-332.
71. Bretthauer, R. K. & Castellino, F. J. (1999). Glycosylation of Pichia pastoris-derived proteins. *Biotechnol. Appl. Biochem.* **30**, 193-200.
72. Montesino, R., Garcia, R., Quintero, O. & Cremata, J. A. (1998). Variation in N-linked oligosaccharide structures on heterologous proteins secreted by the methylotrophic yeast Pichia pastoris. *Protein Expr. Purif.* **14**, 197-207.
73. Jacobs, P. P., Geysens, S., Vervecken, W., Contreras, R. & Callewaert, N. (2009). Engineering complex-type N-glycosylation in Pichia pastoris using GlycoSwitch technology. *Nat. Protoc.* **4**, 58-70.
74. Vervecken, W., Callewaert, N., Kaigorodov, V., Geysens, S. & Contreras, R. (2007). Modification of the N-glycosylation pathway to produce homogeneous, human-like glycans using GlycoSwitch plasmids. *Methods Mol. Biol.* **389**, 119-138.
75. Potgieter, T. I., Cukan, M., Drummond, J. E., Houston-Cummings, N. R., Jiang, Y., Li, F., Lynaugh, H., Mallem, M., McKelvey, T. W., Mitchell, T., Nylen, A., Rittenhour, A., Stadheim, T. A., Zha, D. & d'Anjou, M. (2009). Production of monoclonal antibodies by glycoengineered Pichia pastoris. *J. Biotechnol.* **139**, 318-325.
76. Li, H., Sethuraman, N., Stadheim, T. A., Zha, D., Prinz, B., Ballew, N., Bobrowicz, P., Choi, B. K., Cook, W. J., Cukan, M., Houston-Cummings, N. R., Davidson, R., Gong, B., Hamilton, S. R., Hoopes, J. P., Jiang, Y., Kim, N., Mansfield, R., Nett, J. H., Rios, S., Strawbridge, R., Wildt, S. & Gerngross, T. U. (2006). Optimization of humanized IgGs in glycoengineered Pichia pastoris. *Nat. Biotechnol.* **24**, 210-215.
77. Cregg, J. M., Cereghino, J. L., Shi, J. & Higgins, D. R. (2000). Recombinant protein expression in Pichia pastoris. *Mol. Biotechnol.* **16**, 23-52.
78. Higgins, D. R. (2001). Overview of protein expression in Pichia pastoris. *Curr Protoc Protein Sci*, Unit5 7.
79. Couderc, R. & Baratti, J. (1980). Oxidation of methanol by the yeast, Pichia pastoris. Purification and properties of the alcohol oxidase. *Agric. Biol. Chem.* **44**, 2279-2289.
80. Digan, M. E., Tschopp, J., Grinna, L., Lair, S. V., Craig, W. S., Velicelebi, G., Siegel, R., G.R., D. & Thill, G. P. (1988). Secretion of heterologous proteins from the methylotrophic yeast, Pichia pastoris. In *Development in Industrial Microbiology* (Pierce, G., ed.), Vol. 29, pp. 59-65. Elsevier Science, Amsterdam.
81. Ridder, R., Schmitz, R., Legay, F. & Gram, H. (1995). Generation of rabbit monoclonal antibody fragments from a combinatorial phage display library and their production in the yeast Pichia pastoris. *Biotechnology. (N. Y.)* **13**, 255-260.
82. Nett, J. H. (2009). Production of Antibodies in Pichia pastoris. In *Therapeutic Monoclonal Antibodies: From Bench to Clinic* (An, Z., ed.), pp. 569-584. John Wiley & Sons, Inc., Hoboken, NJ, USA.
83. Takahashi, K., Yuuki, T., Takai, T., Ra, C., Okumura, K., Yokota, T. & Okumura, Y. (2000). Production of humanized Fab fragment against human high affinity IgE receptor in Pichia pastoris. *Biosci. Biotechnol. Biochem.* **64**, 2138-2144.

84. Gach, J. S., Maurer, M., Hahn, R., Gasser, B., Mattanovich, D., Katinger, H. & Kunert, R. (2007). High level expression of a promising anti-idiotypic antibody fragment vaccine against HIV-1 in Pichia pastoris. *J. Biotechnol.* **128**, 735-746.
85. Lange, S., Schmitt, J. & Schmid, R. D. (2001). High-yield expression of the recombinant, atrazine-specific Fab fragment K411B by the methylotrophic yeast Pichia pastoris. *J Immunol Methods* **255**, 103-114.
86. Ogunjimi, A. A., Chandler, J. M., Gooding, C. M., Recinos, A. & Choudary, P. V. (1999). High-level secretory expression of immunologically active intact antibody from the yeast Pichia pastoris. *Biotechnol. Lett.* **21**, 561-567.
87. Lee, B. & Vasmatzis, G. (1997). Stabilization of protein structures. *Curr Opin Biotechnol* **8**, 423-428.
88. Ewert, S., Honegger, A. & Plückthun, A. (2004). Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering. *Methods* **34**, 184-199.
89. Honegger, A., Malebranche, A. D., Röthlisberger, D. & Plückthun, A. (2009). The influence of the framework core residues on the biophysical properties of immunoglobulin heavy chain variable domains. *Protein Engineering Design and Selection* **22**, 121-134.
90. Knappik, A. & Plückthun, A. (1995). Engineered turns of a recombinant antibody improve its in vivo folding. *Protein Eng.* **8**, 81-89.
91. Ewert, S., Honegger, A. & Plückthun, A. (2003). Structure-based improvement of the biophysical properties of immunoglobulin V_H domains with a generalizable approach. *Biochemistry* **42**, 1517-1528.
92. Wörn, A., Auf der Maur, A., Escher, D., Honegger, A., Barberis, A. & Plückthun, A. (2000). Correlation between in vitro stability and in vivo performance of anti-GCN4 intrabodies as cytoplasmic inhibitors. *J Biol Chem* **275**, 2795-2803.
93. Nieba, L., Honegger, A., Krebber, C. & Plückthun, A. (1997). Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. *Protein Eng.* **10**, 435-444.
94. Kügler, M., Stein, C., Schwenkert, M., Saul, D., Vockentanz, L., Huber, T., Wetzel, S. K., Scholz, O., Plückthun, A., Honegger, A. & Fey, G. H. (2009). Stabilization and humanization of a single-chain Fv antibody fragment specific for human lymphocyte antigen CD19 by designed point mutations and CDR-grafting onto a human framework. *Protein Engineering Design and Selection* **22**, 135-147.
95. Jung, S. & Plückthun, A. (1997). Improving in vivo folding and stability of a single-chain Fv antibody fragment by loop grafting. *Protein Eng* **10**, 959-966.
96. Willuda, J., Honegger, A., Waibel, R., Schubiger, P. A., Stahel, R., Zangemeister-Wittke, U. & Plückthun, A. (1999). High thermal stability is essential for tumor targeting of antibody fragments: Engineering of a humanized anti-epithelial glycoprotein-2 (epithelial cell adhesion molecule) single-chain Fv fragment. *Cancer Res.* **59**, 5758-5767.
97. Schimmele, B. & Plückthun, A. (2005). Engineering Proteins for Stability and Efficient Folding. In *Protein Folding Handbook* (Buchner, J. & Kiefhaber, T., eds.), Vol. 5, pp. 1281-1333. Wiley Verlag GmbH & Co. KGaA, Weinheim, Germany.
98. Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J. & Plückthun, A. (2001). Tailoring *in vitro* evolution for protein affinity or stability. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 75-80.
99. Jung, S., Honegger, A. & Plückthun, A. (1999). Selection for improved protein stability by phage display. *J Mol Biol* **294**, 163-180.
100. Fromant, M., Blanquet, S. & Plateau, P. (1995). Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal Biochem* **224**, 347-353.
101. Low, N. M., Holliger, P. H. & Winter, G. (1996). Mimicking somatic hypermutation: affinity maturation of antibodies displayed on bacteriophage using a bacterial mutator strain. *J Mol Biol* **260**, 359-368.
102. Stemmer, W. P. (1994). Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389-391.
103. Moroney, S. & Plückthun, A. (2005). Modern antibody technology: The impact on drug development. In *Modern Biopharmaceuticals* (Knäblein, J., ed.), Vol. 3, pp. 1147-1186. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.
104. Demarest, S. J., Chen, G., Kimmel, B. E., Gustafson, D., Wu, J., Salbato, J., Poland, J., Elia, M., Tan, X., Wong, K., Short, J. & Hansen, G. (2006). Engineering stability into Escherichia coli secreted Fabs leads to increased functional expression. *Protein Eng Des Sel* **19**, 325-336.
105. Monsellier, E. & Bedouelle, H. (2006). Improving the stability of an antibody variable fragment by a combination of knowledge-based approaches: validation and mechanisms. *J Mol Biol* **362**, 580-593.

106. Ewert, S., Cambillau, C., Conrath, K. & Plückthun, A. (2002). Biophysical properties of camelid V_{HH} domains compared to those of human V_H3 domains. *Biochemistry* **41**, 3628-3636.
107. Ewert, S., Huber, T., Honegger, A. & Plückthun, A. (2003). Biophysical properties of human antibody variable domains. *J Mol Biol* **325**, 531-553.
108. Demarest, S. J. & Glaser, S. M. (2008). Antibody therapeutics, antibody engineering, and the merits of protein stability. *Curr. Opin. Drug Discov. Devel.* **11**, 675-687.
109. Shire, S. J. (2009). Formulation and manufacturability of biologics. *Curr. Opin. Biotechnol.* **20**, 708-714.
110. Di Paolo, C., Willuda, J., Kubetzko, S., Lauffer, I., Tschudi, D., Waibel, R., Plückthun, A., Stahel, R. A. & Zangemeister-Wittke, U. (2003). A recombinant immunotoxin derived from a humanized epithelial cell adhesion molecule-specific single-chain antibody fragment has potent and selective antitumor activity. *Clin. Cancer Res.* **9**, 2837-2848.
111. Braun, A., Kwee, L., Labow, M. A. & Alsenz, J. (1997). Protein aggregates seem to play a key role among the parameters influencing the antigenicity of interferon alpha (IFN-alpha) in normal and transgenic mice. *Pharm. Res.* **14**, 1472-1478.
112. Hermeling, S., Crommelin, D. J., Schellekens, H. & Jiskoot, W. (2004). Structure-immunogenicity relationships of therapeutic proteins. *Pharm. Res.* **21**, 897-903.
113. Schellekens, H. (2005). Factors influencing the immunogenicity of therapeutic proteins. *Nephrol. Dial. Transplant.* **20 Suppl 6**, vi3-9.
114. Rosenberg, A. S. (2006). Effects of protein aggregates: an immunologic perspective. *AAPS J.* **8**, E501-507.
115. Kessler, M., Goldsmith, D. & Schellekens, H. (2006). Immunogenicity of biopharmaceuticals. *Nephrol. Dial. Transplant.* **21 Suppl 5**, v9-12.
116. Maas, C., Hermeling, S., Bouma, B., Jiskoot, W. & Gebbink, M. F. (2007). A role for protein misfolding in immunogenicity of biopharmaceuticals. *J. Biol. Chem.* **282**, 2229-2236.

3. Results

3.1 Transfer of engineered biophysical properties between different antibody formats and expression systems

Jonas V. Schaefer and Andreas Plückthun

in **Protein Engineering, Design, and Selection** (2012), Antibody Special Issue

Abstract

Recombinant antibodies and their derivatives are receiving ever increasing attention for many applications. Nonetheless, they differ widely in biophysical properties, from stable monomers to metastable aggregation-prone mixtures. Previous work from our laboratory presented the combination of structure-based analysis with family consensus alignments as being able to improve the properties of immunoglobulin variable domains. We had identified a series of mutations in the variable domains that greatly influenced both the stability and the expression level of scFv fragments produced in the periplasm of *E. coli*. We now investigated whether these effects are transferable to Fab fragments and IgG produced in bacteria, *Pichia pastoris* and mammalian cells. Taken together, our data indicate that engineered mutations can increase functional expression levels only for periplasmic expression in prokaryotes. In contrast, stability against thermal and denaturant-induced unfolding is improved by the same mutations in all formats tested, including scFv, Fab and IgG, independent of the expression system. The mutations in V_H also influenced the structural homogeneity of full-length IgG, and the reducibility of the distant C_H1-C_L inter-chain disulfide bond. These results confirm the potential of structure-based protein engineering in the context of full-length IgGs and the transferability of stability improvements discovered with smaller antibody fragments.

Keywords: IgG, antibody engineering, *Pichia pastoris*, protein stability, denaturation

Abbreviations used:

ADCC, Antibody-dependent cellular cytotoxicity; CD, Circular dichroism; CDC, Complement-dependent cytotoxicity; CDR, Complementarity determining region; CHO, Chinese hamster ovary cells; CIEX, Cation exchange chromatography; DMEM, Dulbecco's modified Eagle medium; DMSO, Dimethyl sulfoxide; DSC, Differential scanning calorimetry; DSF, Differential scanning fluorimetry; DTT, Dithiothreitol; ELISA, Enzyme-linked immunosorbent assay; FBS, Fetal bovine serum; GdnHCl, Guanidine hydrochloride; HEK cells, Human embryonic kidney cells; HPLC, High-performance liquid chromatography; IEF, Isoelectric focusing; IgG, Immunoglobulin G; ITF, Intrinsic tryptophan fluorescence; MRE, Mean residue ellipticity; MS, Mass spectrometry; MST, Microscale thermophoresis; PBS, Phosphate buffered saline; pI, isoelectric point; POI, Protein of interest; PP, *Pichia pastoris;* SEC, Size-exclusion chromatography; SDS-PAGE, Sodium dodecyl sulfate-polyacrylamide gel electrophoresis; WT, wild type

Introduction

Since the emergence of murine monoclonal antibodies as diagnostic and therapeutic agents in the late 1970s, the possibility to generate recombinant human and engineered antibodies by many technologies has led to their greatly accelerated adoption. The reliable and inexpensive production of various antibody formats in recombinant form has remained one of the major objectives for antibody engineering, as antibodies differ widely in their biophysical properties. Small antibody species like single-chain Fv (scFv) or Fab expressed in *Escherichia coli*[1-3] have been pivotal research intermediates in essentially all antibody engineering projects, but by far most antibodies used in the clinic today have been converted to the immunoglobulin G (IgG) format.[4-6] The main reason is that most clinical applications rely upon the effector function of the IgG Fc region. Even if the expression of full-length IgG has been performed successfully in *E. coli*,[7-9] therapeutic antibodies generally are produced in mammalian cells, since only this expression system can carry out the posttranslational modifications and introduce the complex glycosylation necessary for most of the functional activities of IgG molecules.

Compared to conventional small molecules, antibody drugs offer advantages concerning their extended half-life *in vivo* and their ability for initiating host immune responses such as antibody-dependent cellular cytotoxicity (ADCC) or complement-dependent cytotoxicity (CDC).[10] The long half-life of IgG molecules is due to the pH-dependent binding of the C_H2-C_H3 interface to the neonatal receptor FcRn, a major histocompatibility complex (MHC) class I-like molecule expressed on the vascular endothelium. Binding to FcRn protects antibodies from degradation by enabling recycling of the molecules and thereby increases their serum persistence.[11,12]

Residues located in the hinge region and the C_H2 domain contact the Fcγ receptors, and among them especially the activating FcγRIIIa is responsible for triggering ADCC by recruiting immune cells that lead to phagocytosis or lysis of the targeted cells.[10,13-16] The latter interaction requires the complex glycosylation provided by mammalian expression systems, which is not provided by yeast and plant expression systems.[17,18] In CDC, antibodies activate the complement cascade at the cell surface, upon being triggered by C1q binding to the C_H2 domain — eventually killing the targeted cells.[19]

IgGs are composed of various domains, all of them possessing the characteristic immunoglobulin fold consisting of two antiparallel β-sheets forming a β-sandwich,[20,21] which can aggregate, probably involving the β-strands, as in many aggregates of other proteins.[22] The overall stability of the full-length IgG molecule depends on both the intrinsic stabilities of the individual domains as well as on the stability of the corresponding interfaces.[23-26] The various IgG domains are known to feature intrinsic differences in their conformational stability and therefore might also be responsible to different degrees for the biophysical properties of the antibody as a whole.

The IgG's antigen binding sites are formed by the complementarity determining regions (CDRs) within the N-terminal variable regions of both the light chain (V_L) and heavy chain (V_H) coming together and generating the unique surface that specifically recognizes and binds the corresponding antigens.[27]

The remaining framework of the variable domains is quite conserved in both sequence and structure. Nonetheless, different families of variable domains with varying framework sequences have shown quite distinct biophysical properties (see below). The particular arrangement of the antigen binding region within the CDRs thus permits re-engineering of the framework residues without jeopardizing antigen binding, at least to a first approximation.

Due to the great sequence diversity within their variable regions, antibodies show a wide range of biophysical properties.[28-30] Favorable properties are important for a number of reasons.[31] First, an aggregation-prone molecule would run the danger of leading to unspecific binding *in vivo*, premature clearance, loss of clinical efficacy and the danger of eliciting a T-cell-independent immune response.[32,33] Second, the aggregation-tendency would be potentiated in fusion proteins with additional aggregation-prone domains or multimerized constructs, rendering many antibody fusions impossible to be produced efficiently. Finally, low stability and a tendency to aggregate will decrease the expression level, and may thus make otherwise promising molecules economically not viable.

In previous work,[28] we discovered surprising differences in the biophysical properties of the germline families of human V_H domains, as well as both human domains encoding the κ and λ light chain ($V_κ$ and $V_λ$, respectively). By sequence comparison of the human consensus V_H domains, particular mutations could be identified (Fig. 3.1.1a) that proved to be responsible for the favorable biophysical properties of families 1, 3 and 5, as well as for the less favorable properties of families 2, 4 and 6. Subsequently, these sequence differences were transplanted from a V_H3 consensus framework to the very poorly stable and poorly expressing V_H6 framework.[34] Briefly, two non-glycine residues with positive φ angles in the Ramachandran plot were converted to glycines (positions 16 and 76; numbering scheme according to Honegger and Plückthun,[35] inducing a positive effect on both thermodynamic stability and functional expression in *E. coli.* Since functional periplasmic expression levels are limited by folding,[28] they constitute an excellent measure of aggregation during *in vivo* folding. In addition, two residues with non-optimal β-strand forming features were exchanged to amino acids possessing higher β-sheet propensity (Q5V and S90Y), both affecting functional expression, but only the first having a significant thermodynamic equilibrium effect. The removal of a buried non-hydrogen-bonded OH group (T58I) led to a very large increase in thermodynamic stability, but no effect on folding yield. Conversely, the mutation of a solvent-exposed hydrophobic residue to a hydrophilic one (V72D) increased the *in vivo* folding yield but had very little effect on thermodynamic stability.

All the selected mutations (Q5V, S16G, T58I, V72D, S76G and S90Y) were located within the V_H framework and not in the CDR regions, as these mutations were intended to be as generally applicable as possible and unlikely to affect antigen binding. As intended, the combined mutations conferred properties almost identical to the favorable V_H3 framework on the V_H6 framework,[34] which afterwards experienced a dramatic increase in both expression yield and thermodynamic stability without having altered binding properties.

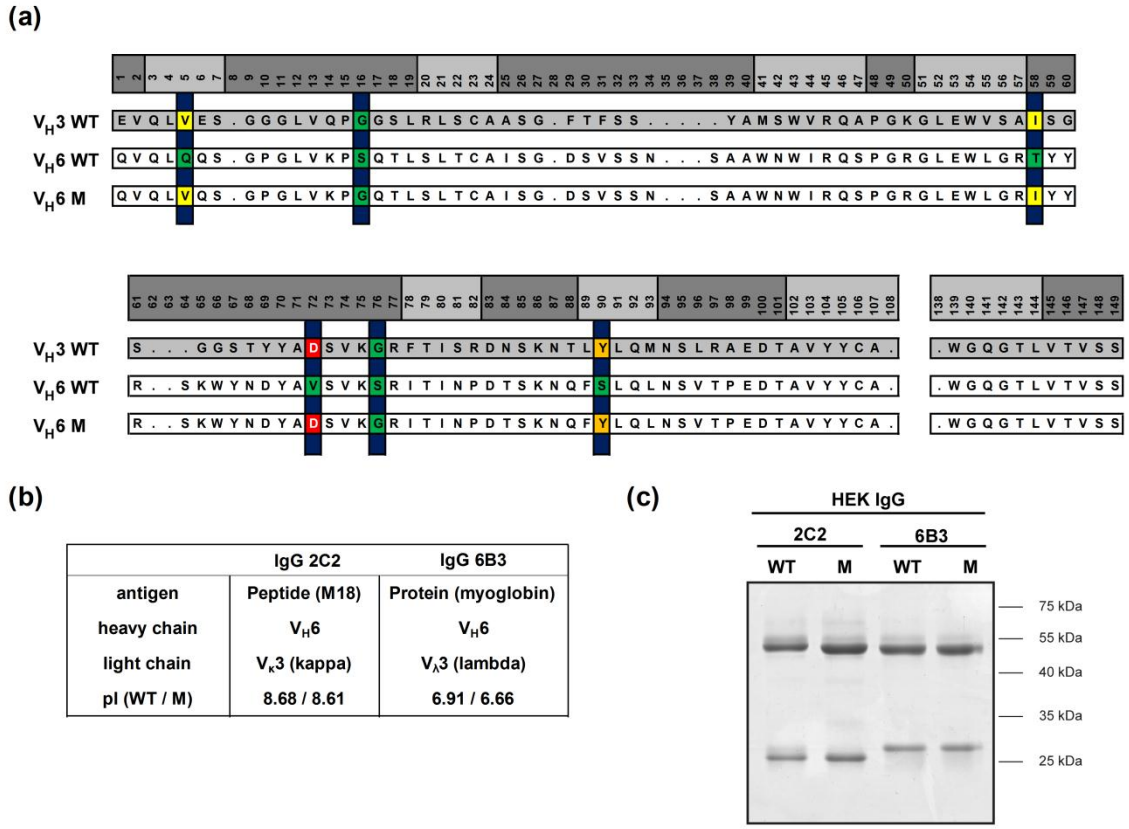


Fig. 3.1.1. Alignment of V_H sequences. (a) Comparison of HuCAL-derived wild-type (WT) V_H3 and V_H6 sequences (two upper most sequences). The V_H6 domain, having aggregation-prone behavior and the lowest midpoint of denaturation, was brought closer to the V_H3 sequence as described before,[34] leading to the engineered M variant (lower line). This mutant holds the following six amino acid exchanges: Q5V, S16G, T58I, V72D, S76G and S90Y. Only mutated amino acids are highlighted and color-coded according to residue type: aromatic (orange), hydrophobic (yellow), uncharged hydrophilic residues (green) and acidic (red). The numbering scheme is according to Honegger and Plückthun.[35] **(b)** Summary of the two different IgG constructs 2C2 and 6B3 analyzed in this study. **(c)** SDS-PAGE analysis of equal amounts of the Protein A-purified IgG constructs under reducing conditions stained with Coomassie Blue.

These data were recorded for two model antibody scFvs named 2C2 and 6B3 that had originally been selected from the HuCAL library[36] by panning against the peptide M18 coupled to transferrin or myoglobin from horse skeletal muscle, respectively. We chose these molecules as both are representatives of the human V_H6 framework family whose members are known to have a rather aggregation-prone behavior and the lowest thermodynamic stability of all human V_H domains.[28] 2C2 and 6B3 have different CDR-H3 loops, while otherwise their V_H sequence is identical. In addition, the two antibodies (compared in Fig. 3.1.1b) differ in their respective light chain, with 2C2 having a $V_κ3$ and 6B3 containing a $V_λ3$ light chain. Importantly, the detected improvements were independent of both the V_L domain as well as the sequence and length of the CDR-H3, as both constructs gave similar results.[34]

These improvements were uncovered and investigated with antibody constructs small enough to untangle the contributions of the individual domains with spectroscopic investigations, using scFv constructs which consist of just two domains. However, the question arose, whether the rather

dramatic effects of these mutations might somehow be dampened in a larger assembly, such as a whole IgG, and whether the eukaryotic secretory quality control[37] might overcome all folding issues seen in *E. coli.* Thus, we now tested the influence of these well-characterized mutations in the context of full-length human IgG1 or Fab molecules expressed in eukaryotic systems. Both the wild-type (WT) and best engineered variant (carrying all six mutations described above, called "M" for "mutant" in the present study) were converted to the Fab or full-length IgG format and investigated for their biophysical properties and expression characteristics upon production in HEK293 cells (HEK) and the yeast *Pichia pastoris* (PP).

To be able to carry out such studies with whole IgGs or Fab fragments, eukaryotic expression systems were required where functional expression yields would exclusively depend on the protein sequence. Therefore, these systems must rely both on homologous recombination to exactly the same locus and on constitutive expression to eliminate complication from the induction strategy. For Fab and IgG production in HEK293 cells, stable cell lines were created using the Flp-In system (Invitrogen). These cell lines expressed and secreted the antibody constructs under control of the constitutive cytomegalovirus (CMV) promoter in sufficient amounts, allowing the purification and subsequent biophysical analysis of these proteins. As *P. pastoris* does not maintain episomal vectors, expression cassettes were anyway designed for integration into the yeast genome. To ensure a comparable expression setup, the constitutive GAP promoter was used.

In the present study, we compare the effects of the previously described mutations on the biophysical properties of antibodies in the IgG and Fab format, produced in mammalian and yeast cells, using a wide array of biophysical techniques. We also analyze the effect of the mutations on the expression levels in bacteria, yeast and mammalian cells — allowing a rough differentiation of how the various cellular hosts deal with aggregation-prone proteins in general.

Results

Expression levels of engineered Fab and IgG formats

We first analyzed the secretion levels of correctly assembled antibodies in the Fab and IgG format by ELISA (depicted in Fig. 3.1.2). They show the expected behavior in reducing SDS gels (Fig. 3.1.1c). To compare the improvements caused by the mutations, the expression yields are normalized relative to the yield of the corresponding WT variants. While for the scFv fragments produced in *E. coli* the engineered M variant previously showed an increase in yield of 4.2- or 4.3-fold,[34] respectively (Fig. 3.1.2a), the variable domain mutations only had a slight influence on the expression level of antibodies in eukaryotic expression systems. While for the *Pichia* system (Fig. 3.1.2d), the M variants still were secreted at 15 - 40% increased levels, no difference could be detected for the expression of either Fab fragments or full-length IgGs in the HEK system (Fig. 3.1.2b and 3.1.2c).

To ensure that these results were not distorted by molecules still trapped within the cell due to an overload of the secretory pathway, the secreted IgG amounts were compared to the intracellular levels (Supplementary Data Fig. S3.1.1a). Considering that the secreted samples of each construct were diluted 1:6 compared to the corresponding intracellular fractions, the blot unambiguously indicated that the majority of molecules were secreted for both IgG variants. Similar results were obtained when analyzing the intra- and extracellular IgG levels in the *Pichia* system (data not shown). Interestingly, the analysis of IgG levels secreted to the periplasm of *E. coli* had quite a different outcome (Supplementary Data Fig. S3.1.1b): the comparison by western blot clearly showed increased expression levels for the M variants even in the IgG format in the prokaryotic system.

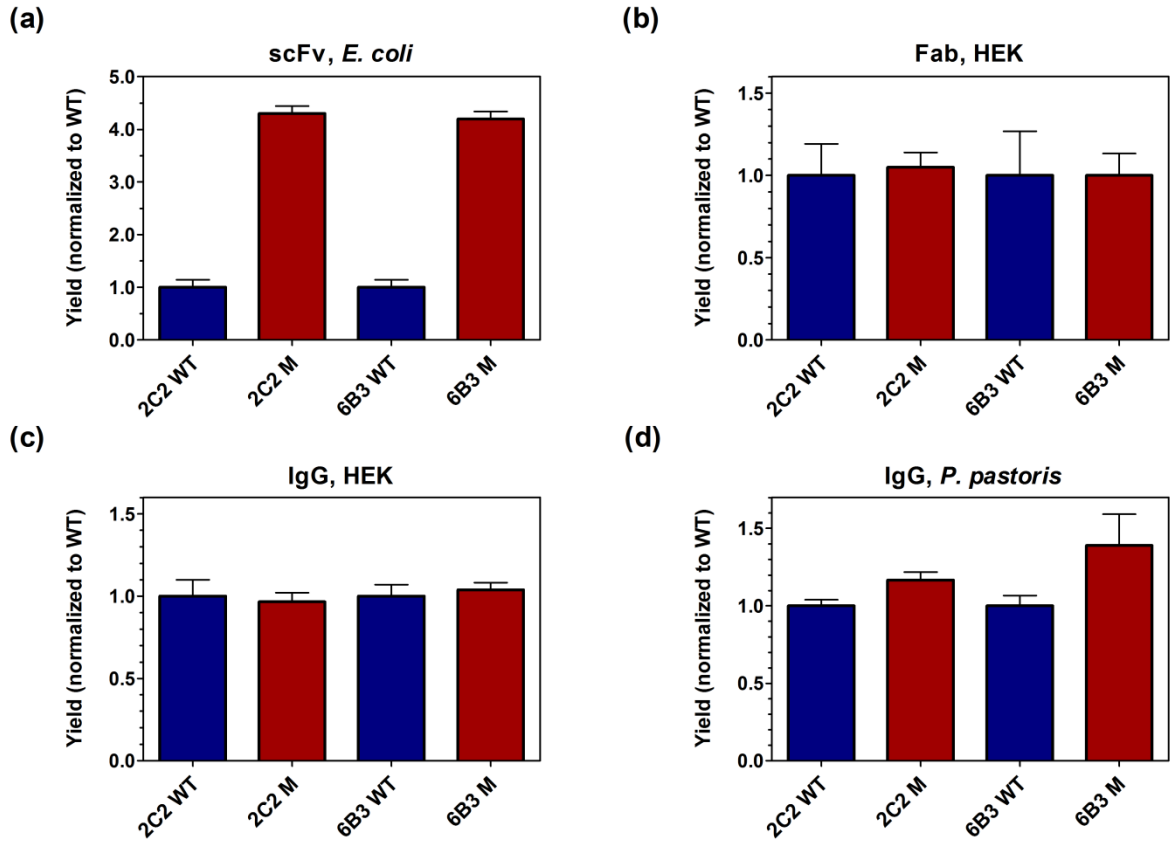


Fig. 3.1.2. Influence of mutations on secretion levels of antibody constructs. Secreted antibody levels of either scFv, Fab or IgG constructs detected by ELISA. For a better comparison, the secretion yield of soluble protein is normalized relative to the yield of the corresponding WT construct. The WT constructs are displayed in blue and the M variants in red. **(a)** Levels of scFv fragments found in the periplasm of *E. coli*.[34] **(b)** Amounts of Fab fragments secreted by stable HEK clones. **(c)** IgG levels found in the supernatant of stable HEK clones. **(d)** IgG amounts detected in the supernatant of stable *Pichia* clones.

Analysis of biophysical properties of full-length IgG molecules

For an analysis of their biophysical properties, IgGs were expressed either in the HEK or PP system in large scale and purified by Protein A affinity chromatography. The samples were of high purity and gave rise to both expected bands under reducing conditions (see Fig. 3.1.1c; the heavy chain was detected at ~ 50 kDa and the respective light chains in the range of 23 - 27 kDa). As depicted in

Supplementary Data Fig. S3.1.2 in extracts for the HEK produced IgGs, the purified WT and M molecules behaved as monomers in size exclusion chromatography (SEC) but showed different elution profiles in reverse phase HPLC (RP-HPLC) and cation exchange chromatography (CIEX). These differences, however, could be accounted for by the introduced mutations themselves, influencing both the hydrophobicity as well as the charge of the engineered M variants. The net hydrophobicity of the engineered M variants was slightly increased by the two mutations Q5V and T58I, while only being lowered by a single mutation (V72D). The analysis of reduced samples by RP-HPLC (depicted in the lower halves of Supplementary Data Fig. S3.1.2c and S3.1.2d) clearly indicated that the difference in the running behavior was caused by the mutated heavy chains, as the corresponding light chains — eluting earlier than the heavy chains due to their lower hydrophobicity — led to identical signals for both variants. The additional negative charge through the V72D mutation influenced the CIEX running behavior of the engineered IgG, causing the M variants to elute at higher volumes (Supplementary Data Fig. S3.1.2e and S3.1.2f).

Finally, the binding affinity of the IgG molecules towards their corresponding antigens was confirmed to be fully retained (Supplementary Data Fig. S3.1.3). From microscale thermophoresis measurements,[38] the K_D for the HEK IgG 2C2 was calculated to be 217 ± 54 nM and 270 ± 76 nM for the WT and M variants, respectively, while the analogous variants for the 6B3 constructs showed affinities of 15.7 ± 4.8 nM and 14.5 ± 3.8 nM, respectively. The slight differences of these K_D values are all within the experimental error range.

To confirm that the mutations did not influence the overall structure of the new antibody molecules, HEK and PP produced IgGs were analyzed by circular dichroism (CD). As shown in Fig. 3.1.3a and 3.1.3b, the new M constructs still possessed the β-sheet structure typical for IgG molecules with only slight alterations. Since the signals stemming from the IgG's β-sheets and random coil essentially canceled out to zero at 208 nm, thermal denaturation of the IgG molecules was determined at this wavelength as a function of temperature (Fig. 3.1.3c - f). The abrupt upward jumps seen between 75°C and 90°C in most of the depicted CD vs. temperature curves were caused by the formation of insoluble aggregates in the respective samples. Interestingly, the used expression system (HEK vs. PP) not only caused a slightly different temperature of unfolding onset, but, unexpectedly, also resulted in very different aggregation susceptibilities of the IgGs — with the 6B3 variants expressed in *Pichia* not exhibiting any detectable aggregation at all (Fig. 3.1.3f). These effects, caused by differences in the N-linked glycosylation and by an incomplete processing of the α-factor pre-pro sequence at the yeast constructs, have been studied extensively and reported in chapter 3.2 and elsewhere.[39] For the present study, however, we concentrated on the comparison of WT and M variants. For the 2C2 construct, no significant difference in stability between WT and M could be detected for IgGs produced in either expression system, and the different signal level in Fig. 3.1.3c for the M variant was caused by its slightly off-set curve as also depicted in Fig. 3.1.3a. We currently do not know what causes this offset. For the 6B3 M construct, however, increased stabilities of about 2.5°C were seen compared to 6B3 WT expressed in either the HEK or the PP system, as shown in Fig. 3.1.3d and 3.1.3f.

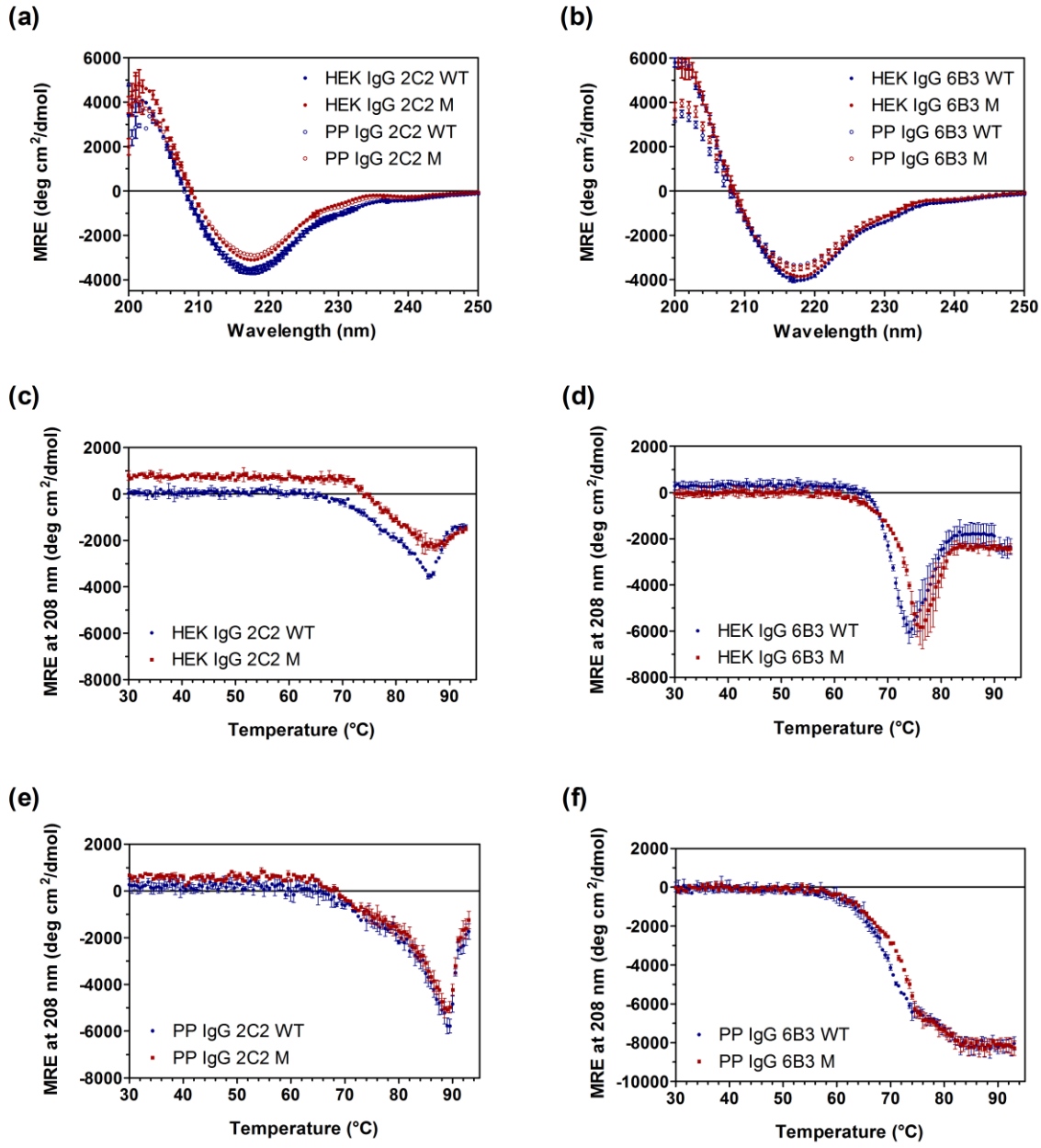


Fig. 3.1.3. Analysis of IgG constructs by circular dichroism (CD). Signals derived from WT variants are displayed in blue while those of the mutant M are represented in red. **(a)** CD spectra of different IgG 2C2 variants produced in mammalian cells (HEK) or in the yeast *Pichia pastoris* (PP), respectively. The values are reported as mean residue ellipticity (MRE). **(b)** CD spectra of different IgG 6B3 variants analogous to panel (a). **(c)** Thermal denaturation curves of IgG 2C2 produced in mammalian cells. The denaturation was followed by CD, plotting the signals at 208 nm as a function of temperature. The values are reported as mean residue ellipticity (MRE). The abrupt upward jump at about 85°C in the signals is caused by the formation of insoluble aggregates. **(d)** Thermal denaturation curves of HEK IgG 6B3 analog to panel (c). The onset temperature of aggregation for this construct was found to be lowered to about 75°C. **(e)** Thermal denaturation curves of *Pichia*-produced IgG 2C2 variants. **(f)** Thermal denaturation curves of *Pichia*-produced IgG 6B3 variants.

Since the method of analyzing thermal stabilities by CD averages all changes observed over the whole IgG molecule and thus does not allow to attribute them to particular domains, samples were further analyzed by intrinsic tryptophan fluorescence (ITF) as a function of temperature. Since ~70% of the tryptophan residues of the studied IgGs are localized in the Fab fragments, this method should be

more sensitive to small changes in the stability of these parts of the molecule. As shown in Fig. 3.1.4a and 3.1.4b, the data derived from CD spectroscopy could be confirmed by ITF, validating the increased stability of the M variants. A clear difference in stabilities could be seen for the 6B3 constructs (Fig. 3.1.4b), where the IgG and Fab WT variants exhibited their transition at 67.6° ± 0.1°C and 69.7° ± 0.1°C and thus at lower temperatures compared to their engineered M counterparts with transition at 70.8° ± 0.1°C and 74.2° ± 0.2°C, respectively. For the 2C2 constructs, the advantages of the M variants, however, were considerably less pronounced (Fig. 3.1.4a).

To probe whether differences in the temperature-induced unfolding of 2C2 might have been masked by inter- or intramolecular aggregation, the same IgG samples were analyzed in the presence of guanidine hydrochloride (GdnHCl). After an overnight incubation in 1 M GdnHCl before being exposed to increasing temperature still in the presence of 1 M GdnHCl, both IgG variants were clearly destabilized and started to unfold at lower temperatures (Fig. 3.1.4c) — however, the M variant showed a slightly increased stability of the second transition at 71.8° ± 0.2°C, compared to its WT counterpart with 70.4° ± 0.3°C.

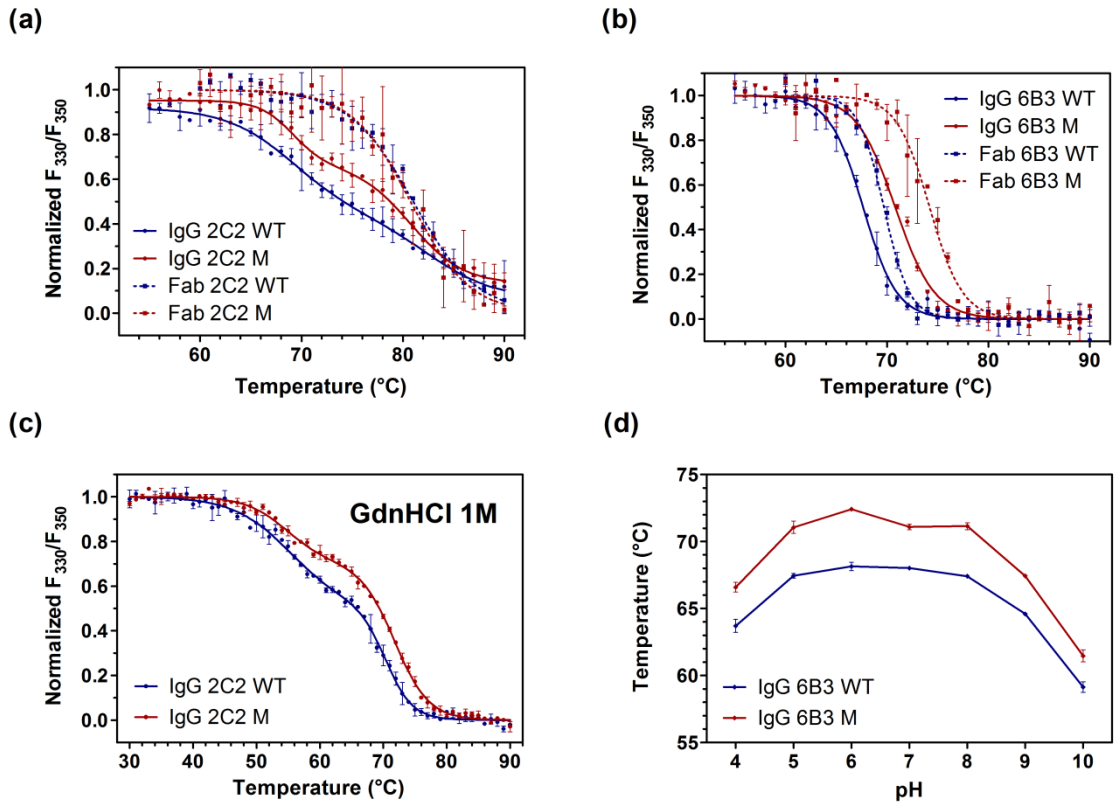


Fig. 3.1.4. Biophysical characterization of IgG and Fab constructs produced in mammalian cells by intrinsic tryptophan fluorescence (ITF). (a) Thermal denaturation curves measured by ITF. The curves were obtained from the intensity ratio of the emission spectrum at 330 nm (F_{330}) and 350 nm (F_{350}) upon excitation at 295 nm, plotted as a function of temperature. A comparison between the different mammalian IgG 2C2 (continuous lines) and Fab 2C2 (dotted lines) variants is shown. **(b)** Thermal denaturation curves of IgG 6B3 and Fab 6B3 measured analogous to the method used in panel (b). **(c)** ITF analysis of HEK IgG 2C2 in the presence of 1 M guanidine hydrochloride (GdnHCl). **(d)** pH dependence of thermal denaturation. The midpoints of the denaturation curves for HEK IgG 6B3 WT and M were determined at various pH and plotted as a function of pH.

To finally challenge the stability of the engineered variant with yet another parameter, the thermal stabilities of the 6B3 WT and M molecules were determined over a broad pH range. As shown in Fig. 3.1.4d, the increased stability of the engineered IgG proved true in all tested buffer systems having pH values between 4 and 10. On average, the stabilities of the M variants were 3.3° ± 0.6°C higher than those of the corresponding WT molecules.

The increased stability of the M variant over its WT counterpart was further investigated with respect to denaturant-induced unfolding. For this purpose, antibody molecules were incubated in various concentrations of GdnHCl and analyzed by ITF after an overnight incubation (Fig. 3.1.5). To better identify the signals derived from unfolding of the C_H2 domains, glycan knock-out T299A mutants were also examined (data not shown). Previous studies[39] presented in chapter 3.2 had clearly shown a decreased C_H2 stability for unglycosylated IgGs, and thus, the first transition e.g. at 2 M GdnHCl in Fig. 3.1.5a could be assigned to C_H2, as it shifted to even lower GdnHCl in the glycan knock-out mutant.

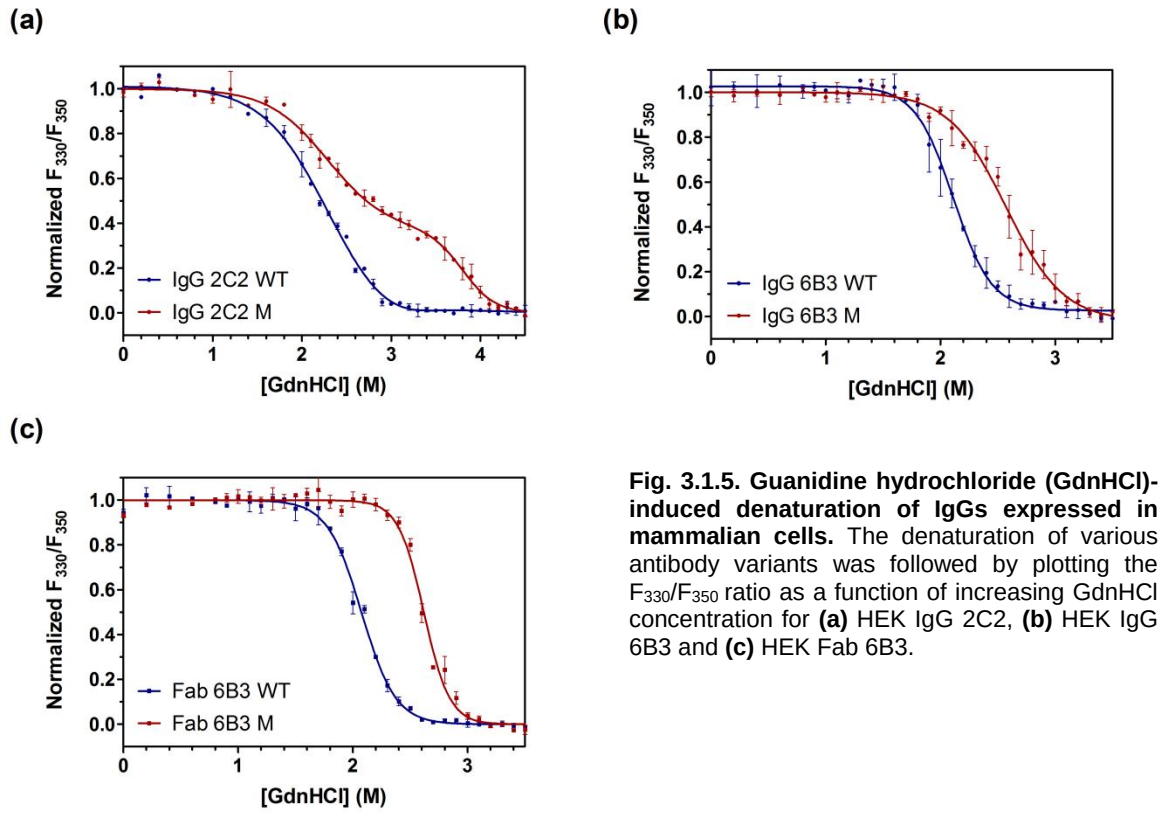


Fig. 3.1.5. Guanidine hydrochloride (GdnHCl)-induced denaturation of IgGs expressed in mammalian cells. The denaturation of various antibody variants was followed by plotting the F_{330}/F_{350} ratio as a function of increasing GdnHCl concentration for **(a)** HEK IgG 2C2, **(b)** HEK IgG 6B3 and **(c)** HEK Fab 6B3.

For all analyzed antibody variants, the introduced mutations caused an increased stability compared to the respective wild type. The second transition of the unfolding curves could be assigned as belonging to the Fab fragment. Whereas this transition was detected in the WT variants of the 6B3 and 2C2 constructs at 2.0 and 2.5 M GdnHCl, the engineered mutants showed an increased Fab stability of 2.6 and 3.8 M, respectively. The same results were also obtained when the Fab 6B3 fragments were analyzed by themselves, i.e. as a produced Fab fragment (Fig. 3.1.5c). The kinetics of unfolding was

considerably different between M and WT for some conditions and thus amplified the stability differences. As illustrated in time-dependent unfolding studies displayed in Supplementary Data Fig. S3.1.4, the presence of e.g. 4.5 M GdnHCl was sufficient to denature the 2C2 WT variant completely within ~10 min while IgG 2C2 M was not fully unfolded in the same GdnHCl concentration even after 2 h. Analogous results were obtained for the 6B3 constructs in the presence of e.g. 3.5 M of denaturant.

Analysis of the antibody molecules on the domain level

Finally, we aimed to investigate the effect of the V_H mutations within the antibody at the level of individual domains or interacting folding units. The methods of choice for these studies are differential scanning calorimetry (DSC) or differential scanning fluorimetry (DSF). DSF has the advantage of small sample requirements and easy sample preparation, making it suitable for high-throughput applications and initial screens. However, the nature of transitions are much more difficult to assign, and changes in later transitions (i.e. more stable domains) are more difficult to detect. For IgGs, the transition of the C_H2 unfolding detected at 72°C emerged to be the most dominant signal, making the detection of later transitions involving the most relevant part of the molecule, the Fab fragment, quite difficult or even impossible (Fig. 3.1.6).

While the transition of a second domain could still be clearly seen for the 6B3 construct (74.5°C vs. 77°C for WT and M variants, respectively, as depicted in Fig. 3.1.6b), DSF showed apparent limitations in the analysis of Fab fragments with very high stabilities. In the 2C2 variants only slight differences were detectable at ~84°C in Fig. 3.1.6a. For comparison, also the corresponding first derivatives (dF/dt) were plotted in Fig. 3.1.6c and 3.1.6d, highlighting again the differences in the second domain transition for the 6B3 IgGs (73.8°C vs. 76.3°C). These signals could be definitely attributed to the corresponding Fab domains, as they matched the data obtained from the corresponding recombinant Fab fragments (Fig. 3.1.6e and 3.1.6f). For the 6B3 Fab variants, the transitions were determined to be 72.8°C or 77.3°C and thus 4.5°C apart.

As DSF proved to be limited in its analytical potential of domains with much higher stabilities than the first transition, antibody samples were subjected to DSC analysis (Fig. 3.1.7). These analyses require more time and sample, but can resolve three major transitions originating from the C_H2, C_H3 and Fab fragments, as is typically seen for IgG.[29] Yet for most of our samples, some of these peaks were overlapping. However, combining the results from both 2C2 and 6B3 constructs as well as from the T299A glycan knock-out mutants with their lowered C_H2 stability (Fig. 3.1.7e), clear assignments could be made. While the transitions of the glycosylated C_H2 and of the C_H3 domains were detected at 70.1°C or 81.3°C, respectively, the signals of the different WT and M Fab fragments were found between 70°C and 88°C for the 6B3 and 2C2 construct.

In agreement with the results from other analysis techniques detailed above, the Fab fragment of the engineered 6B3 M variant was of higher stability than its WT counterpart (on average 2.5°C). This

could be detected for the IgGs produced in HEK (Fig. 3.1.7b), in *Pichia* (Fig. 3.1.7d) as well as for the HEK-produced Fab fragments (Fig. 3.1.7f). Interestingly, for the first time also small differences in the thermodynamic stability could be demonstrated for the 2C2 Fab construct (Fig. 3.1.7a and 3.1.7c). The signals of the mutated Fab fragments showed a melting temperature which was ~1.8°C higher than that of the WT variants and this difference was found for IgGs expressed in either the HEK or PP system, respectively. A comparison of the DSF and DSC derived data is presented in Supplementary Data Fig. S3.1.5.

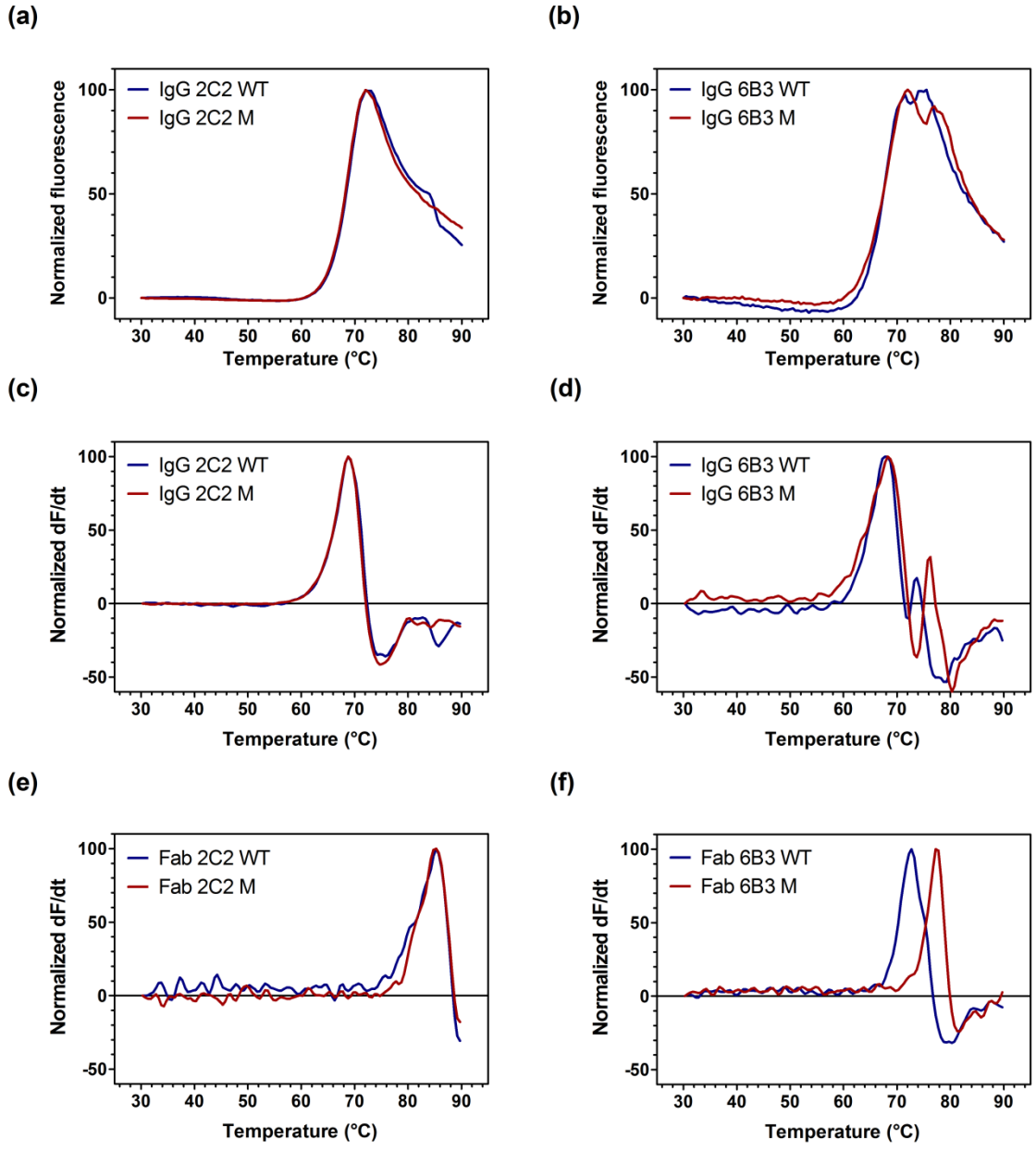


Fig. 3.1.6. Characterization of HEK-produced antibodies by differential scanning fluorimetry (DSF). Comparison of normalized DSF signals of HEK-produced antibody WT (blue) and M (red) variants. For the IgGs, the transition of the C_H2 unfolding is the dominant signal detected. DSF spectra were recorded for **(a)** IgG 2C2 and **(b)** IgG 6B3. For an easier comparison, the corresponding normalized first derivatives (dF/dt) of the signals in (a) and (b) are shown in **(c)** for IgG 2C2 and in **(d)** for IgG 6B3. The dF/dt graphs for the corresponding Fab-fragments are depicted in **(e)** for Fab 2C2 and in **(f)** for Fab 6B3.

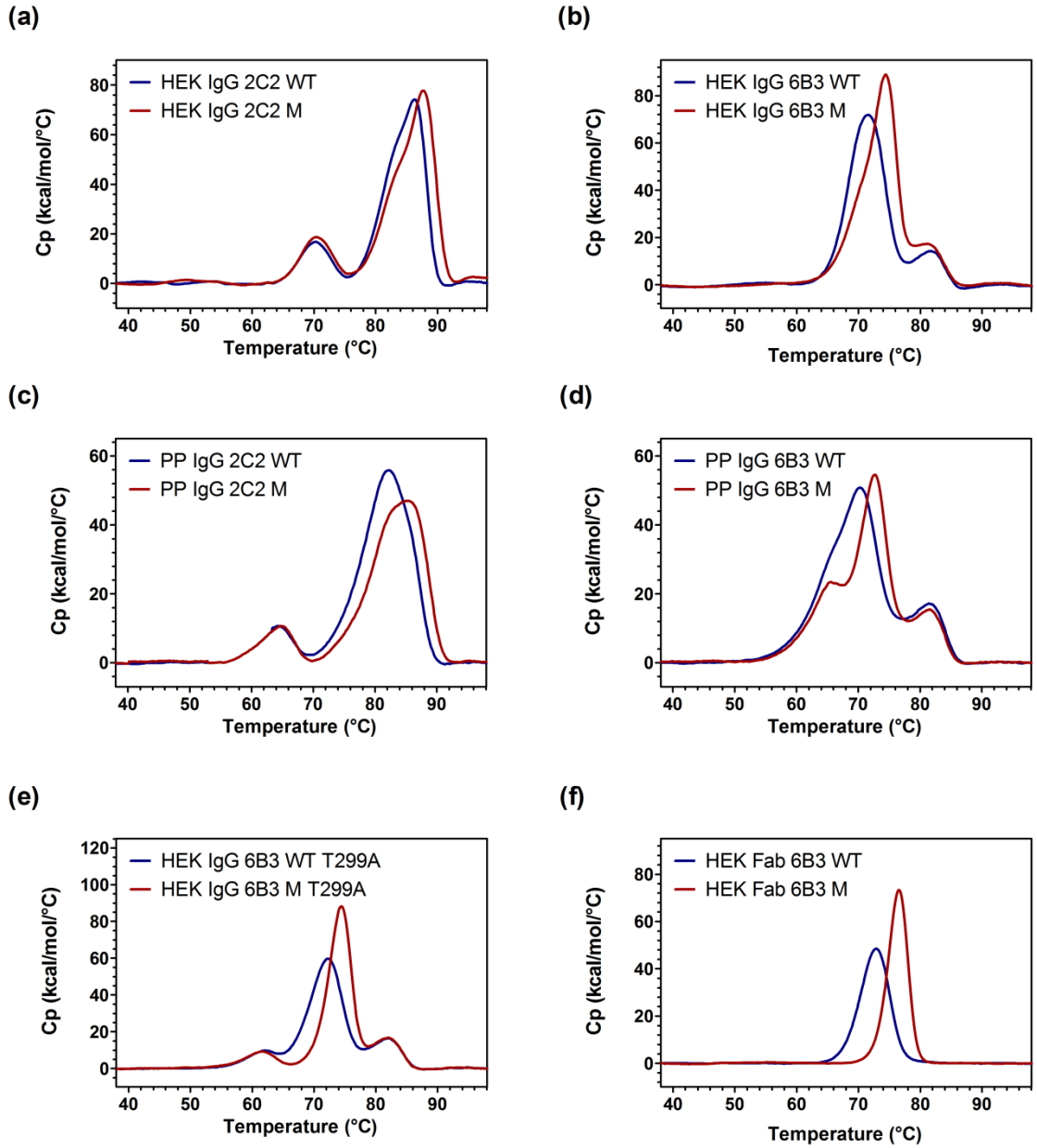


Fig. 3.1.7. Analysis of IgG domain stability by differential scanning calorimetry (DSC). Signals derived from WT variants are shown in blue while those of the mutant M are represented in red. DSC curves were recorded for **(a)** HEK IgG 2C2, **(b)** HEK IgG 6B3, **(c)** PP IgG 2C2, **(d)** PP IgG 6B3, **(e)** HEK IgG 6B3 T299A and **(f)** HEK Fab 6B3.

Electrophoretic analyses of various IgG variants

When full-length IgGs were run on a gradient gel under non-reducing conditions and without prior heating (Fig. 3.1.8a), clear differences for the WT and M variants were encountered. Independent of whether the IgGs had been purified from HEK or PP cultures, WT IgGs usually could not be detected as a clear band at the corresponding size of 150 kDa, but rather as a smear or overlay of several individual bands. In contrast, the M variant gave rise to one clear and distinct band on Coomassie-stained gels (especially the 6B3 constructs, shown on the right side of the gel in Fig. 3.1.8a). To

investigate this peculiar difference in running behavior between WT and M versions of the IgGs in more detail and at greater resolution, the same samples were run under identical conditions on low percentage (7.5%) SDS polyacrylamide gels (Fig. 3.1.8b). Again, the engineered mutants could be detected as one distinct band, while for the WT variant several bands became visible. Interestingly, all these WT bands ran at slightly higher apparent molecular weights than the M counterparts — a feature that could be observed independent of whether the IgGs have been purified from stable HEK or *Pichia* cells.

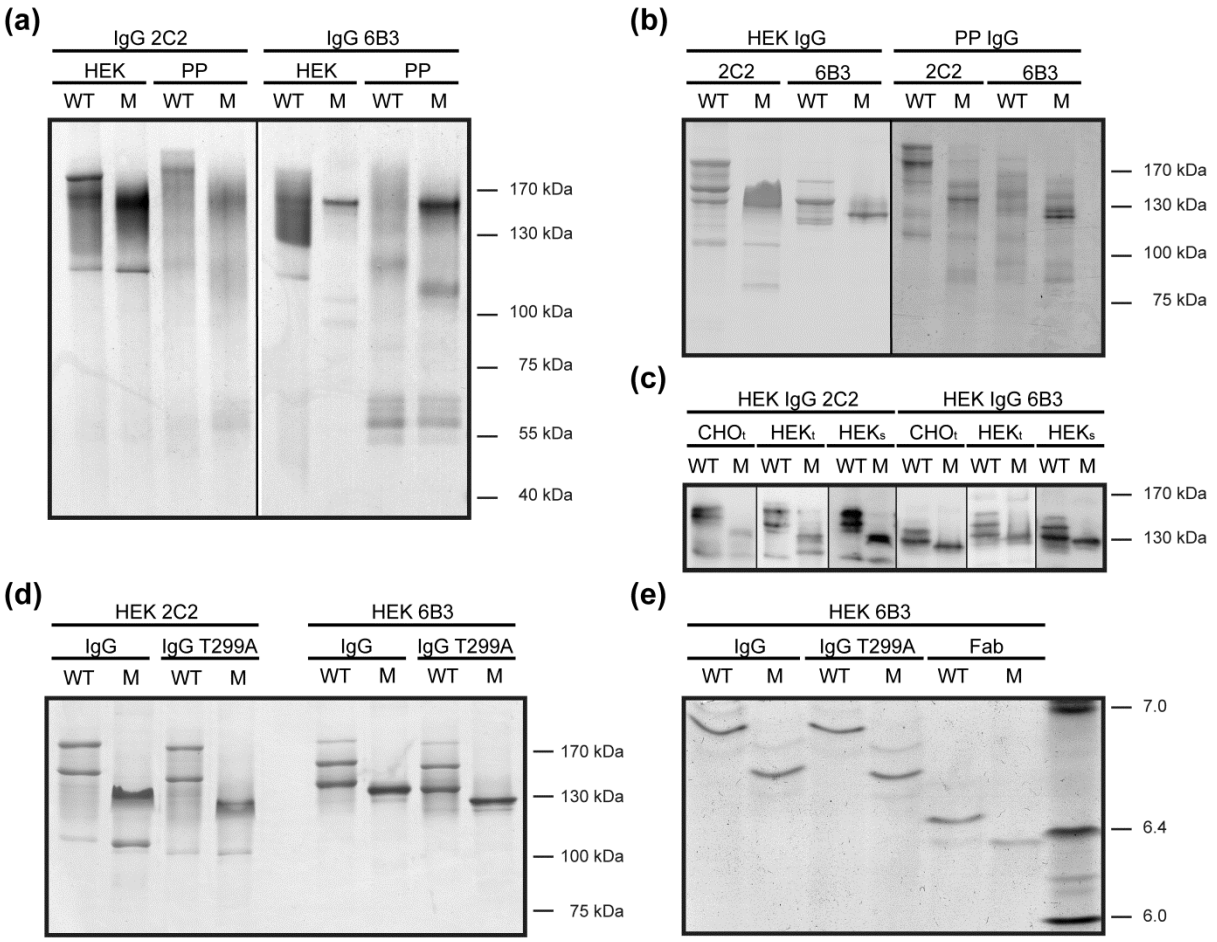


Fig. 3.1.8. Analysis of running behavior of IgGs on SDS-PAGE and IEF. (a) SDS-PAGE of equal amounts of the Protein A-purified HEK and *Pichia* constructs stained with Coomassie Blue. Samples were run under non-reducing conditions without heating prior to loading onto NuPAGE gradient gels. **(b)** SDS-PAGE analysis of the same samples as in panel (a) run on a low-percentage (7.5%) gel under the same, non-reducing conditions. **(c)** Western blot analysis of IgG variants detected with antibodies specific for the respective light chains. Samples were obtained from the supernatant of Chinese hamster ovary cells (CHO) after transient transfection and from HEK293 cells after transient (HEK_t) or stable (HEK_s) transfection, respectively. **(d)** Comparison of HEK IgG variants to their non-glycosylated counterparts (T299A glycan knock-outs) by non-reducing SDS-PAGE. **(e)** Isoelectric focusing of HEK-produced 6B3 variants. Non-reduced IgG, IgG T299A and Fab were separated based on their isoelectric point (pI). A pI standard is shown in the right lane, with the corresponding pI values denoted next to it.

To ensure that these differences between WT and M were not caused by effects of protein purification or storage, we also analyzed the non-processed supernatants of Chinese hamster ovary cells (CHO) after transient transfection and compared it to that of HEK293 cells after transient (HEK_t) or stable (HEK_s) transfection, respectively. As can be seen in Fig. 3.1.8c, the same banding patterns were

observed in all cases and further confirmed for the yeast system as well (data not shown). To ensure that the differences in the running behavior were not caused by variations in the glycan moiety attached to the C_H2 domain of the IgG, antibodies with a glycan knock-out T299A mutation were also analyzed by SDS-PAGE. While, as expected, the unglycosylated IgGs ran generally at slightly lower apparent molecular weights, the banding pattern was not altered (Fig. 3.1.8d).

To test for covalent heterogeneities, we subsequently performed isoelectric focusing (IEF) analyses. In this method, the samples were neither heated nor modified but just analyzed on pH gradient IEF gels. Since the isoelectric points (pI) of the 2C2 constructs were rather high, only the 6B3 constructs could be analyzed by IEF. As shown in Fig. 3.1.8e, all 6B3 constructs (IgG, IgG T299A as well as the Fab fragments) were focused in single major bands at positions of the gel corresponding to their pI, indicating that charge heterogeneity in the molecules was not the reason for the different running behavior on SDS-PAGE. The homogeneity of the samples was also confirmed by mass spectroscopy after prior reduction (data not shown), and is consistent with reducing SDS-PAGE (Fig. 3.1.1c) where no differences in the IgG chains between WT and M had been detected.

To further elucidate the reasons for the different running behaviors of the IgG variants, samples were once more studied by non-reducing SDS-PAGE (Fig. 3.1.9). This time, however, the samples were partly heated for 5 min at 95°C or alkylated prior to loading. While the heating of the sample to 95°C eliminated the different running behavior of the variants (Fig. 3.1.9a), alkylation of samples by three different alkylation reagents did not change these characteristics (Fig. 3.1.9b).

We next considered the possibility that in the WT variants one or more disulfides might not have formed correctly and thus the molecules could have undergone disulfide shuffling, facilitated by sample heating. Previous studies showed that multiple bands for IgG samples on SDS-PAGE in principle can also arise from loss of disulfide-linked chains after disulfide scrambling.[40] Therefore, we wished to elucidate whether the different individual bands seen for the WT variants consisted of distinct covalent isoforms or whether all these bands were in equilibrium. For this reason, IgG 2C2 WT was run on a regular non-reducing 7.5% SDS-gel (after incubation in the absence or presence of alkylation reagent) and afterwards analyzed by a non-reducing second SDS-PAGE in a perpendicular orientation (see Fig. 3.1.9c and 3.1.9d and Materials and Methods for details). The silver-stained bands clearly indicated that the molecules found in each of the various individual bands in the first dimension differentiated again in the second dimension and gave rise to the usual banding pattern, independent of whether the samples had been alkylated beforehand or not. This finding would not be consistent with disulfide shuffling.

To test more rigorously whether the distinct bands of the WT are really in equilibrium, distinct bands of the WT and M variants were cut from the first gel and placed into the pockets of the second one (Fig. 3.1.9e). Again, the previously distinct bands re-divided into the original banding pattern. Due to the increased sensitivity of the silver staining procedure, even small amounts of the higher molecular

weight bands could be detected for the M variant under these conditions. We thus can rule out disulfide shuffling as the source of the multiple bands detected for the WT IgG.

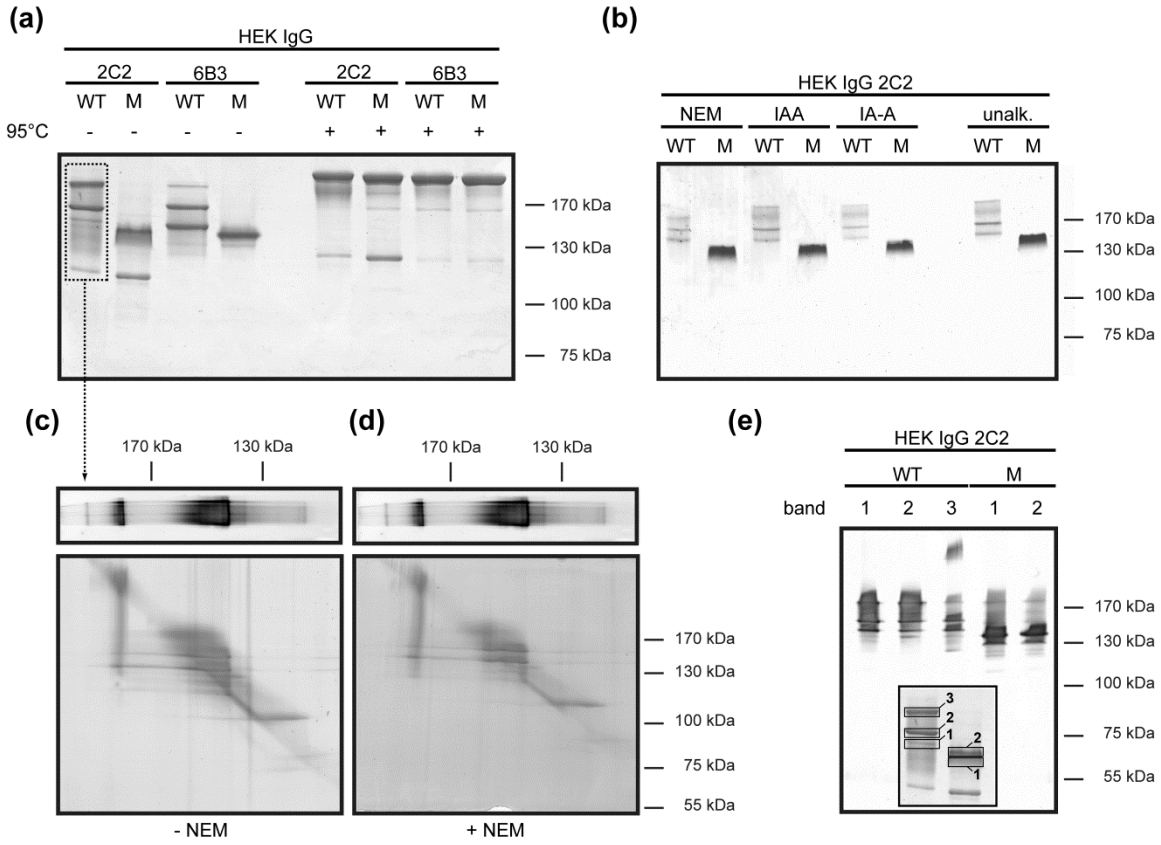


Fig. 3.1.9. One- and two-dimensional SDS-PAGE analysis of the different running behavior of WT and M variants. (a) Analysis of non-reduced IgG samples. Shown are the banding patterns of various IgG constructs after incubation in SDS-loading buffer at either room temperature or 95°C for 5 minutes prior to loading. **(b)** Influence of alkylation on banding patterns. IgG 2C2 samples were incubated with a 50-fold molar excess of N-ethylmaleimide (NEM), iodo-acetamide (IAA) or iodoacetic acid (IA-A) in 100 mM Tris pH 8 (pH 7 for NEM) at 37°C for 3 h prior to analysis by non-reducing SDS-PAGE. **(c)** 2D-SDS-PAGE analysis of HEK IgG 2C2 WT. After the samples had been separated by regular non-reducing SDS-PAGE (shown on top, corresponding to the sample run on lane #1 of the gel depicted in panel (a)), the corresponding lane was isolated and re-run on a second SDS-PAGE in a perpendicular orientation. The final gel was silver-stained. **(d)** 2D-SDS-PAGE of the same sample as in panel (c). However, prior to the first SDS-PAGE, the sample was incubated in the presence of 5 mM NEM for 1 h. **(e)** 2D-SDS-PAGE of HEK IgG 2C2 WT and M. The bands were individually cut out from a first gel, as depicted in the inlay, and run in separate lanes on the second gel (corresponding to the band numbers in the first gel), and the gel was silver-stained.

While both in the present study and the one reported previously[40] multiple bands of IgG have been seen under non-reducing conditions, there are important differences pointing to a different origin of the phenomena. Most importantly, in the experiments reported here the multiple bands *disappear* when the samples are heated to 95°C (Fig. 3.1.9a), while Liu et al. observed an *increase* of the banding pattern (i.e. faster running bands) after prolonged heating and at high pH. Furthermore, in the reequilibration experiments described above (Fig. 3.1.9c-e), all bands can be generated from all other bands (Fig. 3.1.9e), while Liu et al. observed an irreversible increase of smaller bands, consistent with

a loss of chains from the H_2L_2 molecule. We therefore propose that the phenomena seen here are due to conformational rearrangements and *precede* any chain loss due to disulfide scrambling or progressive β-elimination.[40] This is consistent with the western blots, which indicate the presence of both chains in the high MW bands in Fig. 3.1.9, and the generation of a 100-kD band (representing H_2) only after prolonged mild reduction (see below). In summary, the conformational changes seen here are happening under much milder conditions than the chain loss described previously[40] and represent conformational transitions of the intact IgG.

Capillary electrophoresis

To finally exclude that any particular effect of SDS-PAGE itself was causing the multiple band behavior, we also analyzed IgG samples by capillary electrophoresis. In this setup, the detection of the proteins was achieved using SDS micelles containing a provided fluorescent dye.[41] Upon binding to hydrophobic regions on the analyzed proteins, these SDS-dye-complexes are protected from disassembly in the performed dilution step, generating a fluorescent signal clearly above background.

Also in this setup, multiple signals could be detected for the non-reduced, non-heated WT variants, as depicted for the IgG 6B3 WT in lane 1 in Fig. 3.1.10a and for IgG 2C2 WT in Fig. 3.1.10b. Strikingly, the M variants were not detectable at all under these conditions (lane 2 in Fig. 3.1.10a). Only once they had either been heated and/or reduced, equimolar amounts of the loaded M-type became visible in the corresponding chromatograms (Fig. 3.1.10c and 3.1.10d). Thus, the inability to detect the M variant under the initial conditions was not caused by too low IgG amounts loaded onto the chip, but rather hinted at a higher stability or improved folding of this molecule, preventing access to hydrophobic regions by the fluorescent dye used for detection in capillary electrophoresis. Thus, it did not bind to the native M protein, but only to the WT.

Partial reduction of IgG variants indicates differences in conformational stability

As these results indicated differences in the structural stability of the non-reduced, non-heated antibody molecules between their WT and M form, we analyzed HEK-produced IgGs by partial reduction. For this purpose, samples were shortly incubated in the presence of the strongly reducing tris(2-carboxyethyl)phosphine (TCEP) without any further heating prior to loading onto SDS-gels. We reasoned that the hydrophilic nature of TCEP should not allow it to easily penetrate into the hydrophobic cores of the IgG domains where the intradomain disulfides are located, without prior partial unfolding of the domains. Therefore, such partial reduction experiments represent an indicator for the compactness of these domains as well as the packing of the V_H-V_L interface.

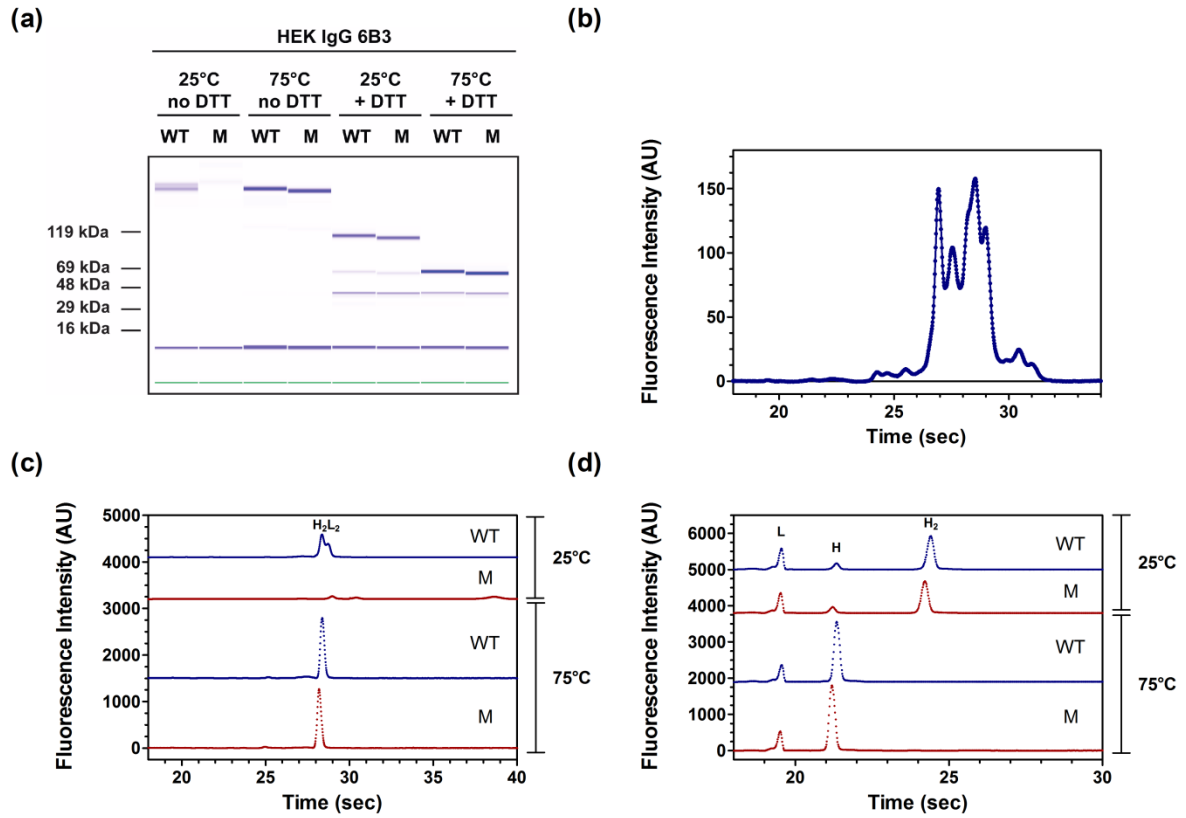


Fig. 3.1.10. Analysis of HEK IgGs by capillary electrophoresis. (a) HEK IgG 6B3 WT and M variants were analyzed in the absence or presence of reducing agent upon incubation in the loading buffer either at 25°C or 75°C (presented as a virtual gel). The same data are presented as chromatograms in panels (c) and (d). **(b)** Chromatogram of HEK IgG 2C2 WT, illustrating the occurrence of similar differences in the banding pattern as detected by SDS-PAGE. **(c)** Chromatograms of IgG 6B3 samples analyzed under non-reducing conditions. H_2L_2 indicates the completely assembled IgG. **(d)** Chromatograms of the same samples as in (c), analyzed after incubation in the presence of DTT. H_2 indicates a heavy chain dimer and H or L the monomeric heavy or light chain molecules, respectively.

As depicted in the Coomassie-stained gel and the corresponding anti-light chain western blot in Fig. 3.1.11a and 3.1.11b, respectively, incubation of the samples in the presence of TCEP seemed to disrupt the disulfide bond between the light and the heavy chain for both the 6B3 and the 2C2 constructs. Interestingly, a much more dramatic effect was observed for the WT constructs than for the M variants. While IgG molecules containing a light chain could still be detected for the 6B3 construct after TCEP reduction, all the HEK IgG 2C2 WT molecules were converted into homo-dimers of the heavy chain only (H_2 molecules), completely lacking the light chains and therefore not being detectable in Fig. 3.1.11b, but only on the Coomassie-stained gel.

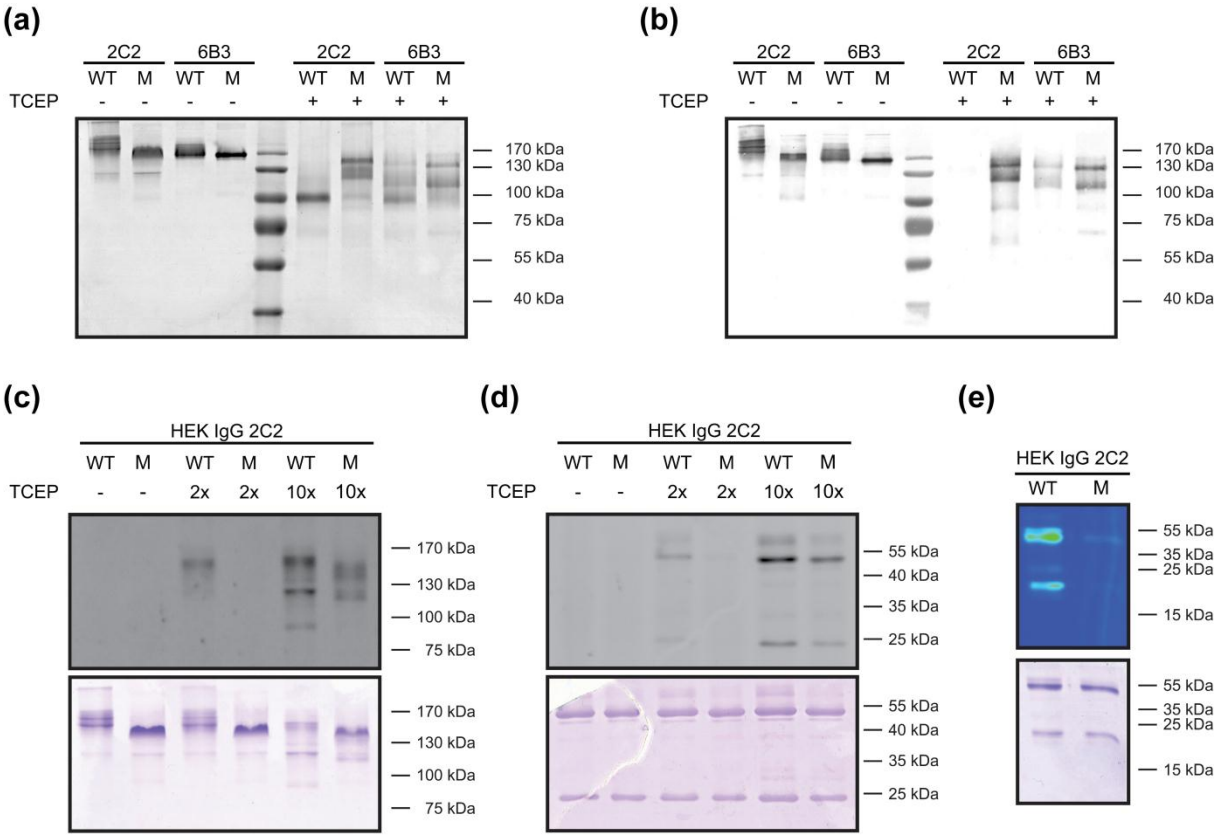


Fig. 3.1.11. Partial reduction of antibody variants. (a) Non-reducing SDS-PAGE of HEK constructs stained with Coomassie Blue. Samples were incubated in the absence or presence of 30 mM TCEP for 10 min at 25°C prior to loading. **(b)** Western blot analysis of the same samples as in panel (a) detected with an antibody mixture specific for both light chains. **(c)** Non-reducing SDS-PAGE of partially reduced HEK IgG 2C2 variants labeled with 5-iodoacetamidofluorescein (5-IAF). Samples were incubated in the absence or presence of the stated molar equivalent of TCEP for 1 h at room temperature prior to labeling with a 5-fold molar excess of 5-IAF for 2 h. The upper gel shows the fluorescence image, the lower picture the same gel after staining with Coomassie Blue. **(d)** Images of the same samples as in (c) run under reducing conditions. **(e)** Comparison of WT and M variants reduced by 2 molar equivalents of TCEP for 1 h prior to labeling with a 5-fold molar excess of 5-IAF overnight. The upper gel shows the fluorescence image of a reducing gel, the lower one the Coomassie-stained gel.

To confirm this observation, partly reduced 2C2 samples were treated with 5-iodoacetamido-fluorescein (5-IAF) to fluorescently label free cysteine residues. Samples were incubated with very low TCEP concentrations, corresponding to only 2 - 10-fold molar equivalents, as higher concentrations were found to interfere with the subsequent 5-IAF coupling. Still, clear differences could be seen in the fluorescent images of non-reducing and reducing gels (Fig. 3.1.11c and 3.1.11d, top gels). Only the WT variant was labeled quantitatively at 2 molar equivalents and to a much higher extent at 10x conditions. Beneath the fluorescent images, the corresponding Coomassie-stained gels confirm that equal antibody amounts were loaded on the gels. Especially after overnight labeling with a 5-fold molar excess (Fig. 3.1.11e), the difference between WT and M variants became clearly visible, as the WT variant was labeled both on its heavy and light chain, while no fluorescence was detectable for the M counterpart.

Taken together, these data suggest that the introduced variable domain mutations, all located within V_H, affected the packing of the V_H-V_L and C_H1-C_L interface and thus protected the intermolecular disulfide bond. It is likely that this improved structural integrity potentially also contributed to the increased stability of the engineered M variants as a whole IgG molecule.

Discussion

The aim of this study was to analyze whether sequence engineering, previously found capable of considerably improving the biophysical properties of antibodies as scFv fragments,[42] would have a similar effect on whole antibodies in the IgG format. The original work was carried out in single domains and smaller assemblies[23; 28; 34; 42-44] as the extremely complex unfolding mechanism of IgGs[45-47] with their 12 immunoglobulin domains interacting with each other would have made an elucidation of individual domain contributions and mutations impossible to untangle in the context of the whole antibody.

Despite the appeal of our previous reductionist approach, it was not clear a priori, whether the previous findings would be transferable from the scFv to the IgG format. It might be possible that the presence of the constant domains of the Fab fragment, especially when linked via an intermolecular C_H1-C_L disulfide bond,[26] overcome stability problems within the variable domains themselves and/or at the V_H/V_L interface. Furthermore, the previous experiments were carried out with antibody fragments produced in functional form by secretion to the periplasm of *E. coli*. Since whole IgGs are predominantly secreted from eukaryotic cells (although functional expression of IgG in *E. coli* has been reported as well;[7-9] it had to be investigated how the eukaryotic quality control system influences both the level and the properties of the resulting molecules, and whether it would be able to somehow compensate for the previously observed differences.

Thus, in this study we aimed to address these questions on two levels: first, the biophysical properties of the resulting proteins in isolation, and second, the influence of the expression systems used. To tackle these questions, we made use of an exceptionally well characterized set of mutants,[34] where prior knowledge for each individual substitution (see Introduction) regarding their effect on thermodynamic stability or folding efficiency was available, measured by functional expression in *E. coli*. As an intermediate between scFv and IgG, we also constructed recombinant Fab fragments. To ensure meaningful results regarding the eukaryotic expression levels, we used homologous recombination into the same chromosomal location of the host cells as well as constitutive promoters for both the *Pichia pastoris* or HEK293 system.

Influence of the expression system

Unlike scFv produced in the bacterial periplasm, IgGs and Fab fragments expressed in both eukaryotic systems did not show any significant differences in their secretion yields found in the respective supernatants (Fig. 3.1.2). We also ensured that almost all of the antibody amounts synthesized by the cells were secreted (Supplementary Data Fig. S3.1.1a), excluding a bottleneck in secretion. Thus, these results derived from HEK and *Pichia* expression indicate that the eukaryotic secretory quality control system equalizes the expression yields of the WT and stabilized V_H6 variants, independent of the antibody format used. Interestingly, the leveling of expression yields is only partial in *Pichia pastoris* (with the stabilized mutants still producing slightly more antibody) but almost complete in the mammalian cells. Similar results had been previously obtained when engineered anti-tetanus toxoid Fab fragments were converted into the IgG format.[48] Also for these Fab molecules the differences in the expression levels seen upon production in *E. coli* were not reflected in different levels of IgG found in mammalian expression.

We could show that the equalized expression levels are due to the eukaryotic host, and not the used IgG or Fab format. When comparing the periplasmic expression levels of full-length IgG in *E. coli*, we found the same increased expression levels for the engineered IgG molecules compared to WT as for the engineered scFv in the same system (Supplementary Data Fig. S3.1.1b). However, since the expression levels of correctly assembled IgG were generally low, no quantitative ELISAs could be performed, and these findings had to be obtained from western blot analysis.

Taken together, all these data indicate that stability engineering can improve bacterial expression and secretion of functional antibodies in all formats tested — single domain scFv, Fab, up to intact IgG — by mutations that prevent aggregation during periplasmic folding.[34; 43; 48-50] However, the eukaryotic quality control can "rescue" such aggregation-prone IgGs and secrete them at comparable level as the engineered mutants.

Analysis of biophysical properties of WT and M

To analyze whether the previously found improvements of the biophysical properties stayed true for various antibody formats produced in eukaryotic systems, sufficient quantities of the corresponding antibodies were purified and analyzed (1-2 mg/L from *Pichia pastoris*, 4-8 mg/L from HEK293 cells), even though the constitutive expression systems were chosen for consistency, and not maximal yields.

Increased stabilities of the engineered IgGs (M variants) were observed by all analysis methods (Figs. 3.1.3 - 7). The data derived from thermal denaturation measured by ITF, DSF and DSC as well as those from GdnHCl unfolding are summarized in Supplementary Data Table S1 for the HEK-produced IgG and Fab fragments. Since the various methods emphasize different factors, the results listed are based on very different phenomena. Therefore, it is useful to briefly summarize what they actually indicate. In general, both CD and ITF measure averaged features of the whole antibody molecule —

either the overall secondary structure or the averaged tryptophan fluorescence, respectively. In contrast, DSF and DSC have the ability, at least in principle, to distinguish different transitions in multidomain proteins (such as IgGs) which are averaged in the previously described spectroscopic methods.

The mid-points of thermal denaturation determined with CD and ITF were in quite good agreement with each other, with the ITF-recorded temperatures being of slightly lower values. Already at these lower temperatures, solvent may start penetrating into the protein core and may thereby make the environment of at least some tryptophans more hydrophilic, while the overall secondary structure measured by CD still is mainly intact. However, the increase in stability determined by both methods for HEK IgG 6B3 M over its WT counterpart gave very similar ΔT results of 2.5° or 3.2°C, respectively. These stability improvements were in good agreement with the data recorded for the Fab fragments by DSF and DSC, even though the measured absolute transition temperatures were of higher value, since the least stable C_H2 domain is not present in the Fab fragment und thus not taken into the average.

For IgGs, DSC analyses generally resulted in three distinguishable transitions, while in the DSF results often only the transition of the C_H2 unfolding could be clearly assigned (for comparison, see Supplementary Data Fig. S3.1.5). This finding can be explained by the fact that DSF measurements are based on intensity changes of the fluorescence of the dye SYPRO Orange as a function of its environment. Its fluorescence is weak in aqueous and hydrophilic environments, but strongly increases upon binding to hydrophobic patches that become accessible in denaturing proteins.[51] After reaching a maximum, the signal decreases at higher temperatures. This is most likely caused by precipitation and aggregation of the analyzed protein, removing SYPRO Orange binding sites from solution. Therefore, the detection of high-temperature transitions, e.g. that of a very stable Fab within the IgG, is very challenging or even impossible.

In contrast, DSC measures the enthalpy of unfolding[52] due to heat denaturation and thus allows the individual detection of various transitions. The DSC results clearly and directly confirmed the increased stabilities of the engineered IgG 2C2 and 6B3 molecules and could assign the effect, as expected, to the Fab fragment within the IgG.

Interestingly, the gain in thermal stability caused by the insertion of the six point mutations was always higher in the context of the isolated Fab fragments compared to the gain recorded within the full-length IgGs. Not just the ITF data revealed a change of 4.5°C compared to only 3.2°C for the IgG molecule, but also the DSC data, allowing to clearly point to a Fab transition, showed an increase of 4°C for the stabilized isolated Fab fragment and 2.2°C for the Fab fragment in the context of the M IgG (Fig. 3.1.7f). At the moment, we have no definite explanation for this phenomenon, and want to stress that these data are derived from irreversible unfolding experiments, where the formation of aggregates could potentially affect the recorded data. We cannot exclude, however, that Fab unfolding might be

influenced by interactions with the Fc part of the IgG, or that kinetic barriers are different for the Fab in isolation and within the IgG — a hypothesis that has been discussed controversially before.[53-55]

Comparing the antibodies 2C2 and 6B3

The overall stabilities of the 2C2 and 6B3 constructs varied significantly, with the 2C2 construct in its wild-type state being already of quite high overall stability. Both antibodies differ in their light chains, with 2C2 containing a κ-light chain with $V_κ3$. This V_L domain had been shown previously[28] to be of considerably higher stability when expressed individually than members of the $V_λ3$ family, which are present in the 6B3 construct. This difference in stability apparently stayed true in the context of single V_L domain, scFv, Fab and IgG, as shown in this study.

Thus, the two antibody constructs analyzed in this study covered a broad range of stabilities, from rather unstable (6B3 constructs) to extremely stable (2C2 constructs). This allowed us to show not only that our engineering strategy was successful in both cases, but also to evaluate analytical methods for detecting the effect in a stable and unstable IgG molecule. We found that most methods were unsuitable to characterize the stability gain of 2C2 M vs. 2C2 WT. Only for thermal denaturation measured by DSC as well as for GdnHCl-induced unfolding, improved stability characteristics of the 2C2 M molecules could be detected. For thermal unfolding studied by ITF, conclusive data could only be obtained in the presence of GdnHCl. For the 2C2 construct, however, the increase in stability was less pronounced with only ~1.6°C compared to ~2.5°C measured on average for the 6B3 constructs.

In GdnHCl-induced unfolding, the engineered variant of 6B3 denatured at a GdnHCl concentration 0.6 M higher than that of the WT counterpart, while this difference was ~1.3 M for the 2C2 constructs. These data could also be confirmed by the analysis of the glycan knock-out T299A mutants (data not shown) and the individual Fab fragments. This difference in stability gain seems puzzling at first, as in thermal stability (measured by a variety of techniques) the impact of the mutations was more pronounced for the 6B3 constructs. However, we must consider two aspects, one regarding the measurements, the other regarding the denaturant: First, the unfolding of a multi-domain protein followed by ITF is very difficult to untangle in either temperature- or denaturant-induced unfolding, since the transitions are not separated and an intra- or intermolecular association of partially unfolded domains will influence tryptophan fluorescence. Second, GdnHCl may prevent such association of partially unfolded domains. Furthermore, the high ionic strength of concentrated GdnHCl may modulate electrostatic contributions during the unfolding. The two analyzed antibody constructs 2C2 and 6B3 possess rather different pI values and thus differ in their charge at any given pH. For these reasons, thermal unfolding may not necessarily be mirrored in GdnHCl unfolding.[56]

While increased stabilities against thermal and denaturant-induced unfolding were also detected for M variants upon expression in *Pichia pastoris* (Fig. 3.1.3f or Fig. 3.1.7c and 3.1.7d), the melting points of IgGs produced in this system were generally lower than for the same IgG produced in mammalian

cells. These different stabilities could be pinpointed to the C_H2-domains and proven to be caused by different glycan moieties, as described in details in chapter 3.2. and elsewhere.[39]

Inhomogeneity of WT but not M in SDS-PAGE

We noticed a consistent ladder formation of WT IgGs on non-reducing SDS-PAGE, while the mutants engineered for stability showed the expected unique band corresponding to a molecular weight of 150 kDa (Fig. 3.1.8a and 3.1.8b). Interestingly, also some commercially available antibodies like Omnitarg® (Pertuzumab, Genentech) showed similar banding patterns in our analyses (data not shown). Proteolysis could be ruled out as a cause considering the results from reducing SDS-PAGE (Fig. 3.1.1c). Similarly, heterogeneity of the glycan structure could be excluded by both IEF analysis (Fig. 3.1.8e) and the fact that the multiple band pattern was preserved in non-glycosylated T299A mutants (Fig. 3.1.8d).

We thus focused our attention on disulfide heterogeneities and/or scrambling, a problem reported to occur for some IgG4 molecules.[57] Yet, alkylation of the samples prior to loading did not affect the running behavior of the IgG (Fig. 3.1.9b). Thus, if different disulfide bond linkages were the reason for the detected phenotypes, they must have been already present in the purified samples. Such observations had been reported previously for a member of the IgG2 family.[58] To test this possibility, we devised a non-conventional 2D-SDS-PAGE (Fig. 3.1.9c - e), in which non-reducing SDS-PAGE was carried out in both dimensions. In this setup, whole lanes or single bands were cut out and re-applied to a new gel. These experiments showed that the different bands could not be separated from each other but re-equilibrated in the second dimension, independent of whether the samples had been previously alkylated or not. We thus can conclude that the multiple bands are not due to disulfide heterogeneity.

This conclusion was further confirmed both by the analysis of unpaired cysteines within the IgG molecules as well as by MS measurement. Quantification of free thiols using the 4-DPS method described by Hansen and colleagues[59] did not indicate any detectable free sulfhydryl groups for the WT and M variants (data not shown), even though this had been detected in other IgG molecules.[60] In addition, analyzing the disulfide patterns of enzymatically fragmented variants using LC-MS coupled to a combination of collision induced dissociation (CID) and electron-transfer dissociation (ETD) fragmentation[61] did not provide any indications that wrong disulfide connectivities were present in any of the IgG variants (data not shown).

Stability probed by dye binding

Stabilized mutants and WT IgGs could be further distinguished by their accessibility either to a fluorescent dye from SDS micelles or to reducing agents. In capillary electrophoresis, which relies on detection of the proteins with SDS micelles containing a provided fluorescent dye,[41] the IgG WT, but not the M variant could be observed under non-reducing, non-heated conditions (Fig. 3.1.10).

However, after heating and/or reduction both variants were equally well detected. These results could not be merely attributed to the very slightly increased hydrophobicity of the M variant, as it was increased only by one amino acid. Therefore, these results rather indicated that the engineered M-type IgGs are more densely packed and, consequently, less label-containing SDS micelles can bind to or penetrate into their structure.

These capillary electrophoresis experiments have some similarity to DSF analyses. However, while in DSF experiment the IgG is heated and the binding of the SYPRO Orange dye is followed as a function of temperature, capillary electrophoresis is isothermal and thus can be carried out below the denaturation temperature. At this temperature, SDS seems to allow access of the dye only to the WT, but not to the stabilized mutant. This is remarkable, since the DSC experiments showed that the C_H2 domains denatured first, and these domains are identical for WT and M. Thus, the significant differences detected for M and WT variants in the capillary electrophoresis experiment must be due to the dye's accessibility to binding sites within the Fab fragment.

Stability probed by partial reduction

Intermediates in the denaturation and reduction of IgGs have been monitored before, both by SDS-PAGE and by capillary electrophoresis.[62; 63] However, we believe that the present study is the first to compare point mutants in the IgG format using these techniques. In partial reduction experiments with TCEP, clear differences in the accessibility of the inter-chain H-L disulfide bond could be detected. The WT variants, but not their M counterparts were losing their light chains (which were identical in both mutants) either to a large extent or completely, as confirmed by both western blot analysis and fluorescent labeling of the liberated thiol groups. This finding was rather astonishing, since all introduced mutations are located within the V_H domain, while the disulfide bond connects the distal ends of C_H1 and C_L. Thus, these data suggest a weakening of the H/L interface which is felt throughout the whole Fab fragment.

Taken together, these results support our assumption that the performed mutations do not just influence the stability of the resulting IgGs, but also a domain other than the one where the mutations are localized (V_H), thus affecting the overall integrity and structural homogeneity of the IgG molecules. It is well established that IgG domains other than the C_H2 (due to its glycosylation) interact strongly in a lateral fashion, forming V_H-V_L, C_L-C_H1 and C_H3-C_H3 modules, respectively.[26] In contrast to these strong lateral interactions, longitudinal interactions generally are reported to be rather weak or even nonexistent. Nonetheless, the ball and socket joints between the V_H and C_H1 or V_L and C_L domains, respectively, combined with some movement within the V_H-V_L interface provide some structural variation to the Fab fragment[64] that could partially explain the detected results.

Whether, however, structural changes within the Fab fragment can be transmitted to the constant domains within the Fc part has been controversially discussed for a long time.[26; 65-68] Our results do not indicate such interactions in the native molecule, and the observations may be explained without

postulating them. Mutations within the V_H domain could affect the V_H-V_L interface, which by the coupling within the Fab fragment[26] could lead to the observed facilitated reduction of the C_H1-C_L disulfide. Antibodies having lost their light chains may subsequently show a different susceptibility to reducing agents in the hinge region and other domains in the Fc part. As a corollary, we can consider the variable domains as a potentially weak link, and the engineering of this domain can make a decisive improvement in stability — even for the whole IgG molecule, as shown by this study.

Conclusions

Taken together, the present data clearly indicate that the mutations introduced into the V_H6 framework have beneficial effects not just in the scFv context but also in the IgG format. While an increase in expression levels was only detectable for periplasmic expression in *E. coli*, increased stabilities both with respect to thermal and denaturant-induced unfolding of both Fab fragments and full-length IgG were seen, independently of the expression host. These mutations proved to be beneficial independent of the nature of the light chain and could be applied to IgGs of rather distinct characteristics concerning their pI values and Fab stabilities. The mutations also influenced the structural integrity and homogeneity of the engineered IgG molecules independently of any wrong disulfide bond connectivity. Presumably through an improved V_H-V_L packing, the reduction of the intermolecular C_H1-C_L disulfide was greatly slowed down in the stabilized variant. These results confirm the potential of structure-based protein engineering in the context of full-length IgGs and the transferability of improvements discovered in a systematic study of smaller antibody fragments. Furthermore, it extends the engineering of intact antibodies to the variable regions, which has mostly been concentrated on the constant Fc part.

Materials and Methods

Cell culture

Materials and cultivation of mammalian and *Pichia* cells

All media and supplements for mammalian expression were purchased either from Sigma-Aldrich (MO, USA), Invitrogen (CA, USA) or Amimed (BioConcept, Switzerland). The antibiotics Zeocin™ and Hygromycin B were bought from Invitrogen or PAA (Austria), respectively. All solutions used were either delivered sterile or sterilized by filtration through 0.22 µm filters (Millipore, MA, USA). Stably transformed human embryonic kidney (HEK) 293 cells were maintained in Dulbecco's Modified Eagle Medium (DMEM, Sigma-Aldrich; high glucose: 4.5 g/l) supplemented with 10% *v/v* heat-inactivated Fetal Bovine Serum (FBS, Amimed) in a humidified incubator with 5% carbon dioxide at 37°C. Expression of IgGs was carried out in DMEM supplemented with 5% *v/v* FBS (instead of the more commonly used 10%). For secretion of IgGs, derivatives of the vector pcDNA5 (Invitrogen) containing constitutively active CMV promoters upstream of the native IgG signal sequences were used (see below).

For all work with *Pichia pastoris,* the strain SMD1163 (*his4 pep4 prb1*; Invitrogen) was used. All media and supplements for this work were purchased either from Sigma-Aldrich or Invitrogen. All work was done in a sterile laminar flow bench, and yeast growth was performed at 30°C. Selection of clones stably expressing the IgGs was based on Zeocin™ resistance. For secretion of IgGs, derivatives of the

pGAPZαB vector (Invitrogen) containing the constitutively active *P. pastoris* GAP promoter followed by the α-factor pre-pro (αMFpp) region were used. Yeast extract-peptone-dextrose (YPD) medium containing 20 g/l peptone, 10 g/l yeast extract and 20 g/l D-glucose (plus 20 g/l agar for YPD-agar) was used for routine growth and subculturing of *Pichia* cells. Zeocin™ was added to a final concentration of 100 µg/ml. IgG expression was performed in the phosphate-buffered medium with glycerol for yeast (BMGY) (20 g/l peptone, 10 g/l yeast extract, 100 mM potassium phosphate pH 6.0, 1.34% yeast nitrogen base (YNB) without amino acids, 1% *w/v* glycerol, 400 µg/l biotin) in the absence of antibiotic.

Construction of Expression Plasmids

Unless stated otherwise, all molecular biology methods were performed according to standard protocols.[69] All enzymes used for cloning were purchased from New England Biolabs (MA, USA) or Fermentas (Germany). In general, cloning and propagation of all plasmids was carried out in *Escherichia coli* DH5α (Life Technologies, CA, USA), grown at 37°C in low salt LB Broth (10 g/l Bacto-tryptone, 5 g/l yeast extract, 5 g/l NaCl, pH 7.5) containing 25 µg/ml Zeocin™ for the *Pichia* plasmids or in 2YT broth (16 g/l Bacto-tryptone, 10 g/l yeast extract, 5 g/l NaCl, pH 7.5) containing 100 µg/ml Ampicillin (AppliChem, Germany) for mammalian vectors.

The construction of expression vectors is described in detail in the Supplementary Material. Briefly, expression vectors for mammalian HEK293 cells were derived from the pMORPH® vector series (MorphoSys), in which heavy or light chain are separately expressed. For the establishment of stable cell lines using the Flp-In system (Invitrogen), the two expression cassettes for H and L, each retaining its own promoter and poly-A sequence, were combined into a single vector based on the pcDNA5/FRT series (Invitrogen). The creation of Fab expression vectors was performed likewise, in this case however, the V_H/C_H1 chain was fused to a myc-tag and $(his)_6$-tag at its C-terminus for detection and purification purposes.

The *Pichia* expression vectors are based on the vector pGAPZαB (Invitrogen). The sequences for the light chains and the different heavy chain variants were amplified from the pMORPH® vectors and thus placed behind the α-factor pre-pro (αMFpp) sequence and under the control of the constitutive GAP promoter. Before stable integration into the *Pichia pastoris* genome, the two expression cassettes for H and L, each retaining its own promoter and termination sequence, were combined onto one plasmid.

Construction of glycan knock-out mutants T299A

To analyze the influence of the glycan moiety attached to Asn297 in the C_H2 domain of the heavy chains, the glycan knock-out construct T299A was constructed, described in detail in chapter 3.2 and elsewhere.[39] Briefly, ACG coding for the threonine in the glycosylation motif Asn_{297}-Ser_{298}-Thr_{299}, was mutated to alanine-encoding GCG using the QuikChange® site-directed mutagenesis kit from Stratagene (acquired by Agilent, USA) according to the manufacturer's instructions.

Creation of stable cell lines

Flp-In HEK293 cell lines (Invitrogen), stably expressing the IgG of interest and secreting it into the culture medium, were generated according to the instructions of the Flp-In system manual. For this purpose, the expression cassettes for both the heavy and light chain genes were cloned into the pcDNA5/FRT vector and finally inserted at a specific location in the genome by the Flp recombinase within the cell. After selection, ten single clones of each construct were tested for expression of the desired IgG by Western blot analysis of the supernatant. More than 90% of the analyzed clones were positive.

For the *Pichia* constructs, the corresponding plasmids containing both genes for the light and heavy chain under the control of constitutively active GAP promoters within one plasmid were linearized by *Bgl*II and integrated into the yeast genome upon transformation of competent *P. pastoris*.[70] Finally, several colonies per construct were re-plated on fresh YPD-agar plates containing Zeocin™ and subsequently analyzed for IgG expression by Western blot.

Expression of IgG

Large scale expression in the mammalian and the yeast system is described in detail in chapter 3.2 and elsewhere.[39] Prokaryotic expression of IgG was performed in *Escherichia coli* SB536 (WG1 ΔfhuA ΔhhoAB[71]) in small scale. Both chains were placed in a bicistronic ORF within pMORPH®X7 derivatives (MorphoSys;[72] under an inducible lac promoter/operator. Pre-cultures of 5 ml 2YT medium containing 1% glucose were inoculated with a single colony and incubated at 37°C overnight. The expression cultures (10 ml 2YT with 0.1% glucose) were inoculated with the pre-cultures at an OD_{600} of 0.1 and grown at 30°C for 20 h. IgG synthesis was not actively induced by adding inducer but rather driven from the leaky expression once all glucose was consumed. IPTG addition only marginally increased the amount of full-length IgG but lead to more partially assembled molecules (like HL and H_2L intermediates). The next day, the cultures were OD-normalized and molecules in the periplasmic space were extracted by incubation of the cell pellet for 2 h at 4°C in the presence of 200 mM boric acid, 160 mM NaCl, 2 mM EDTA, pH 8.0. Finally, the expression levels of WT and M variants were compared by western blots using anti-light chain-specific antibodies as described below.

Biophysical and biochemical analysis

Purification of IgGs

Antibodies were purified from culture supernatants by Protein A affinity chromatography. For this purpose, the supernatants were loaded onto HiTrap Protein A columns (GE Healthcare, USA) at 4°C at a flow rate of 1 ml/min. Chromatography was performed using an ÄKTA PrimePlus chromatography system (GE Healthcare) at 4°C. After loading, the column was washed with 100 mM sodium phosphate buffer pH 8.0 containing 150 mM NaCl. Elution of IgG was accomplished by using 0.1 M glycine pH 2.7, followed by immediate neutralization of each fraction to pH 7.5 using 1 M Tris, pH 8.0. The concentrations of the sample fractions were determined by UV-spectroscopy at 280 nm with a NanoDrop ND-1000 spectrophotometer (Thermo Scientific), assuming a mass extinction coefficient of 1.37 for a 1 mg/ml solution of IgG. The samples with the highest protein concentration were pooled and dialyzed twice against PBS (Sigma-Aldrich; 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 2.7 mM KCl, 137 mM NaCl, pH 7.1) at 4°C. After dialysis, the samples were filtered through 0.22 µm filters (Millipore) and stored at 4°C.

SDS-PAGE and Western blotting

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed using either NuPAGE precasted 4-12% gradient gels (Invitrogen) according the manufacturer's instructions or self-made Tris-glycine 7.5% or 12% gels for non-reduced or reduced samples, respectively. Gels and samples were prepared according to standard protocols[73] and 1.5 µg of IgG antibody was loaded per lane. Non-reduced samples were only mixed with loading buffer, while for reduced samples dithiothreitol (DTT) was added to a final concentration of 16 mM, followed by incubation for 5 min at 95°C. As molecular mass marker the PageRuler prestained protein ladder (Fermentas) was used. The gels were stained using Coomassie following standard protocols or by silver staining according to the manufacturer's instructions (Invitrogen).

For the specific detection of the IgGs, the proteins were transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore) after their separation by SDS-PAGE. Blotting was performed following the semi-dry method within a Trans-Blot SD cell (BioRad, CA, USA) for 60 minutes at 11 V, using a Tris-glycine buffer (20 mM Tris, 150 mM glycine, 10% methanol). After blocking the membrane in 5% MPBST (5% *w/v* skimmed milk dissolved in PBST (PBS with 0.05% Tween-20) for at least 1 h, the membrane was incubated for another hour in 2.5% MPBST containing the appropriate detection antibody conjugated to alkaline phosphatase. For detection of the light chains, an anti-human lambda-light chain-specific antibody (Sigma-Aldrich, #A2904; 1:1,500 dilution) or an anti-human kappa-light chain-specific antibody (Southern Biotech, #2060-04; 1:3,000 dilution) were used. Bands were detected by the addition of NBT-BCIP solution (nitro-blue tetrazolium chloride with 5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt; Sigma-Aldrich) in 100 mM Tris pH 9.5, 5 mM $MgCl_2$, 100 mM NaCl. The development was terminated when the first bands became clearly visible by rinsing the membrane in ultra high pure water.

Isoelectric focusing

Determination of the isoelectric point of the constructs was carried out in IEF Ready Gels (pH 5-8; Bio-Rad, CA, USA) according to the manufacturer's instructions on a Mini-PROTEAN II vertical cell system (Bio-Rad), using increasing voltages (100 V for 1 h, 250 V for 1 h, and 500 V for 30 min). Next to the samples, an IEF standard (Bio-Rad) was co-electrophoresed. The cathode and anode buffers were 20 mM lysine / 20 mM arginine or 7 mM phosphoric acid, respectively. Following electrophoresis, the proteins were visualized by Coomassie Blue - Crocein Scarlet staining (0.04% w:v Coomassie R-250 and 0.05% w:v Crocein Scarlet 3B in 27% isopropanol / 10 % acetic acid) for 45 min and destaining (40% methanol / 10% acetic acid).

Capillary electrophoresis

Capillary electrophoresis was performed on the chip-based Lab-Chip 90 instrument (Caliper Life Sciences, MA, USA) under non-reducing and reducing conditions following the manufacturer's instructions with slight modifications. Briefly, 5 µl sample containing 2.5 µg IgG was mixed with 35 µl of denaturation solution and incubated either at 25°C or 75°C for 5 min. The denaturation solution was prepared according to the provided protocol either without any reducing agent or by the addition of DTT in the recommended concentration. 70 µl distilled water were added to each sample prior to analysis and the analysis chip was primed according to the manufacturer's instructions. The automated protein analysis generated both a chromatogram (showing migration time) as well as a virtual gel (mimicking a Coomassie stained gel) as output.

2D-electrophoresis with SDS-PAGE in both dimensions

To analyze whether the different banding patterns in WT and M samples consist of species which can be separated or species which re-equilibrate, a 2D-SDS-PAGE was devised. Gels were run in a cold room. Duplicates of 4 µg IgG samples were run for the first dimension on an SDS-PAGE gel on two identical halves. While one half of the gel (containing one of the duplicates each) was used for Coomassie staining of the protein bands, the second half was stored temporarily in SDS-running buffer. After the bands were clearly visible, either the corresponding positions or whole lanes on the un-stained gel half were cut and placed on top of the second SDS polyacrylamide gel in perpendicular orientation. This resulted in the re-separation based on the size of the proteins in the second dimension as well. To prevent air bubbles to influence the running behavior, the placed pieces of gel were mounted with regular stacking gel material poured round the gel pieces. The final gel was stained using the Silver Quest kit (Invitrogen) according to the manufacturer's instructions.

Analysis of IgG and Fab expression levels by enzyme linked immunosorbent assay (ELISA)

To analyze the influence of the performed mutations on the antibody expression and secretion levels, the same number of HEK293 cells stably expressing the corresponding IgG or Fab constructs were seeded in 12-well plates. Upon reaching confluency, the medium was removed and replaced by 1 ml expression medium, containing 5% FBS. 24 h later, the supernatant was analyzed for its IgG or Fab content by ELISA. For the *Pichia*-produced IgGs, OD-normalized aliquots were taken after 24 h expression for ELISA analysis. For the IgG detection ELISA, a capture antibody recognizing the human IgG heavy chain (Jackson ImmunoResearch, PA, USA; #209-005-098) was immobilized on MaxiSorb plates (Nunc) overnight at 4°C. For detection of Fab fragments in the supernatant, an anti-myc capture antibody (Sigma-Aldrich, #M4439; 1:2,000 dilution) was coated on the plates. After 1 h blocking in 5% skimmed milk in PBST, 100 µl of *P. pastoris*- or HEK-derived supernatant (diluted in a range of 1:25 to 1:100 in fresh BMGY or DMEM media) was incubated for 1 h at room temperature. The expressed IgG or Fab molecules were detected by incubating with an anti-human lambda-light chain specific antibody conjugated to alkaline phosphatase (Sigma-Aldrich, #A2904; 1:2,000 dilution) for 1 h and subsequent addition of *p*-nitrophenyl phosphate (Sigma-Aldrich). Absorbance at 405 nm was measured using a Perkin Elmer HTS 7000 Plus plate reader by sampling up to 1 h.

To analyze the levels of IgG still present within the cells, cells stably expressing the IgG constructs were cultured in 12-well plates for three days in duplicates. The supernatant was harvested, the cells washed twice with PBS and finally 250 µl of radioimmuno precipitation assay (RIPA) buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1% NP40 and 0.25% Na-deoxycholate) added to the cells before the plates were placed at -80°C for 2 h. After this freezing-thawing step, the cells were scraped off and the extracts were subsequently analyzed by SDS-PAGE followed by Coomassie staining or by western blotting. To allow the simultaneous detection of IgG in the supernatant and the cellular extracts on the

same gel, only 1/6 of the corresponding supernatant was loaded compared to the cellular extract, and all samples were heated to 95°C for 5 min prior to loading.

Circular dichroism spectroscopy

Circular dichroism (CD) measurements were performed on a Jasco J-810 Spectropolarimeter (Jasco, Japan) equipped with a computer-controlled water bath (refrigerated circulator FS18, Julabo, Germany), using a 0.5 mm cylindrical thermocuvette. CD spectra were recorded from 200 to 250 nm with a data pitch of 0.5 nm, a scan speed of 20 nm/min, a response time of 4 s, and a bandwidth of 2 nm. Measurements were performed at 25°C and each spectrum was recorded three times and averaged. The CD signal was corrected by buffer subtraction and converted to mean residue ellipticity (MRE, θ) using the concentration of the sample determined spectrophotometrically at 280 nm. Heat denaturation curves were obtained by measuring the CD signal at 208 nm at temperatures increasing from 25°C to 90°C (heating rate 1°C/min; response time 4 s; bandwidth 2 nm). Data were collected and processed as described above. CD spectra and denaturation curves of the purified IgGs were measured in PBS (Sigma-Aldrich, pH 7.1) at a protein concentration of 5 µM.

Intrinsic tryptophan fluorescence

Intrinsic tryptophan fluorescence (ITF) was measured with a Jobin-Yvon Fluoromax-4 spectrofluorimeter (Horiba Scientific, NJ, USA) equipped with a Peltier-controlled cuvette holder. The temperature was controlled by an LFI3751 5A digital temperature control instrument (Wavelength Electronics Inc., MT, USA). Upon excitation at 295 nm, Trp emission spectra were recorded from 300 to 400 nm ($\Delta\lambda$ = 1 nm, scan rate 1 nm/s) in 0.5°C steps from 25°C to 90°C. The sample cuvette was equilibrated for 2 min at each temperature to ensure that the desired temperature was reached within the cell. Protein concentrations were 1 µM in every case, and all measurements were performed in PBS (pH 7.1). The intensity of the emission spectrum at 330 nm (F_{330}) and 350 nm (F_{350}) was determined at each temperature, the ratio F_{330}/F_{350} calculated and subsequently plotted as a function of temperature.

For the analysis of the midpoints of denaturation at various pH, IgGs were dialyzed overnight and analyzed at the conditions described above in the following buffers. For pH 4 and 5: 25 mM sodium acetate, 130 mM NaCl; for pH 6, 7 and 8: 25 mM sodium phosphate, 130 mM NaCl; for pH 9 and 10: 25 mM sodium carbonate, 130 mM NaCl.

Guanidine hydrochloride-induced equilibrium unfolding

Guanidine hydrochloride (GdnHCl)-induced denaturation measurements were carried out with protein/GdnHCl mixtures containing a final protein concentration of 1 µM and denaturant concentrations ranging from 0 to 5 M (99.5% purity, Fluka, MO, USA). These mixtures were prepared from a 6 M GdnHCl stock solution (in PBS, pH adjusted to 7.1) and equilibrated overnight at 25°C. Each final concentration of GdnHCl was determined by measuring the refractive index. The intrinsic fluorescence emission spectra were then recorded from 300 to 400 nm with an excitation wavelength of 295 nm in an Infinite M1000 reader (Tecan, NC, USA). Individual GdnHCl blanks were recorded and automatically subtracted from the data. The emission ratio F_{330}/F_{350} was calculated and plotted as a function of GdnHCl concentration. Time-dependent unfolding was recorded with the same setup in a Jobin-Yvon Fluoromax-4 spectrofluorimeter (Horiba Scientific, NJ, USA), recording full spectra every two minutes for 5.5 h, starting at the time of GdnHCl addition without any previous incubation.

Size-exclusion chromatography

SEC experiments were performed using a Superdex 200 PC 3.2/30 column (GE Healthcare). The runs were performed in PBS buffer (Sigma-Aldrich, pH 7.1) at a flow rate of 0.06 ml/min at 25°C on an ÄKTA Micro system (GE Healthcare). Samples of 50 µl containing 6.7 µM IgG were injected and protein elution was monitored at 280 nm. Cytochrome c (12.4 kDa), albumin (66 kDa) and β-amylase (200 kDa) were used as standards to calibrate the column.

RP-HPLC analysis

HPLC analyses were performed according to Dillon and colleagues[74] on an Agilent 1100 HPLC system equipped with a binary pump. Samples were separated on a Zorbax SB300 C8, 3.5 µm at 75°C with a flow rate of 0.25 ml/min. The following setup was used for all runs: Solvent A: 0.1% TFA,

solvent B: 0.09% TFA / 9.9% water / 20% acetonitrile / 70% isopropanol. The column was initially equilibrated with 90% mobile phase A and 10% mobile phase B for 5 min followed by a 2 min step gradient from 10 to 25% B. Elution was achieved with a linear gradient of 25 - 40% B over 18 min. Afterwards a final flush of 90% B was performed for 5 min and the system set back to initial conditions and equilibration for 11 min prior to the next run (total run time = 45 min). Reduced samples were adjusted to 4.2 M GdnHCl / 70 mM Tris / 3.5 mM EDTA, pH 8.6 in the presence of 50 mM DTT and incubated at room temperature for 2 h prior to analysis.

Cation-exchange chromatography

Cation-exchange chromatography was carried out on an Agilent 1100 binary pump LC system equipped with a TSKgel CM-STAT column, 7 µm, 4.6 × 100 mm column (Tosoh Bioscience, Japan). 30 µg of purified IgGs were injected onto the column equilibrated in buffer A (25 mM sodium phosphate pH 6.0). Antibodies were eluted by a linear gradient to 30% buffer B (25 mM sodium phosphate pH 6.0, 1 M sodium chloride) over a 30 min time period with detection at 280 nm. The column temperature and flow rate were maintained at 25°C and 0.5 ml/min, respectively.

Differential scanning calorimetry

Differential scanning calorimetry (DSC) measurements were performed using a VP-Capillary DSC system (Microcal Inc., acquired by GE Healthcare). The antibody concentrations were adjusted to 0.5 mg/ml prior to the measurement. The corresponding buffer was used as a reference. The samples were heated from 15°C to 100°C at a rate of 1°C/min after initial 8 min of equilibration at 15°C. A filtering period of 10 s was used and data were analyzed using Origin 7.0 software (OriginLab1 Corporation, MA, USA). Thermograms were corrected by subtraction of buffer-only scans and then normalized to the molar concentration of the protein. The final excess heat capacity thermogram was obtained by interpolating a cubic baseline in the transition region. The midpoint of a thermal transition temperature (Tm) was obtained by analyzing the data using Origin 7 software provided with the instrument. As all measured transitions are irreversible, all the experimental values reported in this study for melting temperatures have to be regarded as "apparent" values.

Differential scanning fluorimetry

Differential scanning fluorimetry (DSF) was performed using the Rotor-Gene Q real-time PCR cycler (QIAGEN) and fluorescence data were collected using the instrument's HRM channel settings (λ_{ex} 460 nm; λ_{em} 510 nm). The SYPRO Orange dye (Molecular Probes; see Volkova *et al*.[75] for a partial structure) in DMSO was diluted 500-fold from the supplied stock solution into the appropriate buffers just prior to being added to the protein solutions. The samples with a final protein concentration of 3.5 µM in 20 µl reaction mixture were subjected to a temperature ramp from 30°C to 90°C at a heating rate of 1°C/min and at 0.5°C increments with an equilibration time of 30 s at each temperature prior to measurement. The T_m was determined as the temperature corresponding to the maximum value of the first derivative of the fluorescence changes, calculated by the software. When multiple unfolding transitions were observed, only the T_m value of the first transition could be accurately determined, as the transitions at higher temperatures overlap. Prior to the DSF analysis, several IgGs were analyzed by fluorescence spectroscopy in the presence or absence of SYPRO Orange. In agreement with the literature, the dye did not induce any changes in the thermal stability determined by ITF, as long as the final dye concentration was lower than a 1:200 to 1:500-fold dilution of the original reagent.

Microscale thermophoresis measurements

Binding affinities of purified IgGs to their antigen myoglobin (Sigma-Aldrich) or M18-Transferin (Jerini GmbH, Germany) respectively were measured using microscale thermophoresis (MST, NanoTemper, Germany) as described previously.[38] Myoglobin or M18-Transferin was fluorescently labeled according to the manufacturer's instructions with a reactive NT-647 dye using N-hydroxysuccinimide (NHS) ester-chemistry which reacts with primary amines to form dye-protein conjugates. For each analyzed construct, a titration series with constant antigen concentration (20 µM) and varying IgG concentrations between 10^{-11} and 10^{-5} M was prepared in PBS. The mixed samples were equilibrated for 1 h at room temperature and approximately 4 µl of each sample was loaded in the capillary. An infrared laser diode within the Monolith NT.115 instrument (NanoTemper, Germany) was used to increase the temperature by 4 K in the beam center. Throughout the measurement, the fluorescence

inside the capillary was recorded by a CCD-camera and the normalized fluorescence was plotted afterwards against the IgG concentrations. The K_D values were subsequently obtained from fitting the binding curves using Prism 5 (GraphPad, CA, USA).

Partial reduction

Partial reduction of IgGs was performed by incubating the samples in 30 mM tris(2-carboxyethyl)phosphine (TCEP) for 10 min at 25°C. Afterwards, the samples were analyzed without further heating by non-reducing SDS-PAGE and either stained by Coomassie or analyzed by western blot.

Fluorescent labeling of free sulfhydryls with 5-iodoacetamidofluorescein (5-IAF, Thermo Fischer) was performed after incubation of IgG with low µM concentrations of TCEP (corresponding to 2-10 molar equivalents; higher concentrations interfered with the subsequent 5-IAF coupling) for 1 h at room temperature. Afterwards the partially reduced samples were incubated with 5-fold molar excess of 5-IAF for 2 h at room temperature protected from light. Afterwards the samples were separated by non-reducing or reducing SDS-PAGE and the fluorescently labeled bands detected using the LAS3000 chemiluminescence detection system (Fuji, Japan). Finally, all protein bands were visualized by Coomassie staining.

Acknowledgements

The authors want to thank Dr. P. Gimeson (GE Healthcare, Sweden), Dr. T. Müller-Späth (ChromaCon, Switzerland) and Dr. D. Weinfurtner (MorphoSys, Germany) for their help with the DSC, CIEX and Capillary electrophoresis experiments, respectively. We would like to acknowledge Dr. M. Heller (University of Bern, Switzerland) for his MS analyses and active help in their interpretation. We are further grateful to the personnel of the Functional Genomics Center at University of Zurich for their help with MS analyses and RP-HPLC. Our thanks are also due to Dr. S. Duhr (NanoTemper, Germany) for his help in determining affinities by microscale thermophoresis. We further thank Dr. P. Lindner for critically reading the manuscript and for valuable suggestions, and the other members of the Plückthun laboratory for fruitful discussions.

J. Schaefer was recipient of a Kekulé predoctoral fellowship of the German Chemical Industry Association and a member of the Molecular Life Science Ph.D. program. This work was supported by the Schweizerische Nationalfonds (SNF) grant 3100A0-128671/1 (to A.P.).

References

1. Skerra, A. & Plückthun, A. (1988). Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli*. *Science* **240**, 1038-1041.
2. Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R. & Opperman, H. (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 5879-5883.
3. Glockshuber, R., Malia, M., Pfitzinger, I. & Plückthun, A. (1990). A comparison of strategies to stabilize immunoglobulin Fv-fragments. *Biochemistry* **29**, 1362-1367.
4. Plückthun, A. & Moroney, S. E. (2005). Modern antibody technology: The impact on drug development. In *Modern Biopharmaceuticals* (Knäblein, J., ed.), Vol. 3, pp. 1147-1186. 4 vols. Wiley-VCH, Weinheim.

5. Beck, A., Wurch, T., Bailly, C. & Corvaia, N. (2010). Strategies and challenges for the next generation of therapeutic antibodies. *Nat. Rev. Immunol.* **10**, 345-352.
6. An, Z. (2010). Monoclonal antibodies - a proven and rapidly expanding therapeutic modality for human diseases. *Protein Cell* **1**, 319-330.
7. Simmons, L. C., Reilly, D., Klimowski, L., Raju, T. S., Meng, G., Sims, P., Hong, K., Shields, R. L., Damico, L. A., Rancatore, P. & Yansura, D. G. (2002). Expression of full-length immunoglobulins in Escherichia coli: rapid and efficient production of aglycosylated antibodies. *J. Immunol. Methods* **263**, 133-147.
8. Mazor, Y., Van Blarcom, T., Iverson, B. L. & Georgiou, G. (2009). Isolation of full-length IgG antibodies from combinatorial libraries expressed in Escherichia coli. *Methods Mol. Biol.* **525**, 217-239, xiv.
9. Makino, T., Skretas, G., Kang, T. H. & Georgiou, G. (2011). Comprehensive engineering of Escherichia coli for enhanced expression of IgG antibodies. *Metab. Eng.* **13**, 241-251.
10. Jefferis, R., Lund, J. & Pound, J. D. (1998). IgG-Fc-mediated effector functions: molecular definition of interaction sites for effector ligands and the role of glycosylation. *Immunol. Rev.* **163**, 59-76.
11. Ghetie, V. & Ward, E. S. (2000). Multiple roles for the major histocompatibility complex class I- related receptor FcRn. *Annu. Rev. Immunol.* **18**, 739-766.
12. Roopenian, D. C. & Akilesh, S. (2007). FcRn: the neonatal Fc receptor comes of age. *Nat. Rev. Immunol.* **7**, 715-725.
13. Shields, R. L., Namenuk, A. K., Hong, K., Meng, Y. G., Rae, J., Briggs, J., Xie, D., Lai, J., Stadlen, A., Li, B., Fox, J. A. & Presta, L. G. (2001). High resolution mapping of the binding site on human IgG1 for Fc gamma RI, Fc gamma RII, Fc gamma RIII, and FcRn and design of IgG1 variants with improved binding to the Fc gamma R. *J. Biol. Chem.* **276**, 6591-6604.
14. Lazar, G. A., Dang, W., Karki, S., Vafa, O., Peng, J. S., Hyun, L., Chan, C., Chung, H. S., Eivazi, A., Yoder, S. C., Vielmetter, J., Carmichael, D. F., Hayes, R. J. & Dahiyat, B. I. (2006). Engineered antibody Fc variants with enhanced effector function. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4005-4010.
15. Strohl, W. R. (2009). Optimization of Fc-mediated effector functions of monoclonal antibodies. *Curr. Opin. Biotechnol.* **20**, 685-691.
16. Kaneko, E. & Niwa, R. (2011). Optimizing therapeutic antibody function: progress with Fc domain engineering. *BioDrugs* **25**, 1-11.
17. Jefferis, R. (2009). Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action. *Trends Pharmacol. Sci.* **30**, 356-362.
18. Lux, A. & Nimmerjahn, F. (2011). Impact of differential glycosylation on IgG activity. *Adv. Exp. Med. Biol.* **780**, 113-124.
19. Ricklin, D., Hajishengallis, G., Yang, K. & Lambris, J. D. (2010). Complement: a key system for immune surveillance and homeostasis. *Nat. Immunol.* **11**, 785-797.
20. Lesk, A. M. & Chothia, C. (1982). Evolution of proteins formed by beta-sheets. II. The core of the immunoglobulin domains. *J. Mol. Biol.* **160**, 325-342.
21. Padlan, E. A. (1994). Anatomy of the antibody molecule. *Mol. Immunol.* **31**, 169-217.
22. Dobson, C. M. (2003). Protein folding and misfolding. *Nature* **426**, 884-890.
23. Wörn, A. & Plückthun, A. (1998). Mutual stabilization of V_L and V_H in single-chain antibody fragments, investigated with mutants engineered for stability. *Biochemistry* **37**, 13120-13127.
24. Brandts, J. F., Hu, C. Q., Lin, L. N. & Mos, M. T. (1989). A simple model for proteins with interacting domains. Applications to scanning calorimetry data. *Biochemistry* **28**, 8588-8596.
25. Wörn, A. & Plückthun, A. (1999). Different equilibrium stability behavior of scFv fragments: Identification, classification, and improvement by protein engineering. *Biochemistry* **38**, 8739-8750.
26. Röthlisberger, D., Honegger, A. & Plückthun, A. (2005). Domain interactions in the Fab fragment: a comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability. *J. Mol. Biol.* **347**, 773-789.
27. Mian, I. S., Bradwell, A. R. & Olson, A. J. (1991). Structure, function and properties of antibody binding sites. *J. Mol. Biol.* **217**, 133-151.
28. Ewert, S., Huber, T., Honegger, A. & Plückthun, A. (2003). Biophysical properties of human antibody variable domains. *J. Mol. Biol.* **325**, 531-553.
29. Garber, E. & Demarest, S. J. (2007). A broad range of Fab stabilities within a host of therapeutic IgGs. *Biochem. Biophys. Res. Commun.* **355**, 751-757.
30. Wang, N., Smith, W. F., Miller, B. R., Aivazian, D., Lugovskoy, A. A., Reff, M. E., Glaser, S. M., Croner, L. J. & Demarest, S. J. (2009). Conserved amino acid networks involved in antibody variable domain interactions. *Proteins* **76**, 99-114.

31. Demarest, S. J. & Glaser, S. M. (2008). Antibody therapeutics, antibody engineering, and the merits of protein stability. *Curr. Opin. Drug Discov. Devel.* **11**, 675-687.
32. Maas, C., Hermeling, S., Bouma, B., Jiskoot, W. & Gebbink, M. F. (2007). A role for protein misfolding in immunogenicity of biopharmaceuticals. *J. Biol. Chem.* **282**, 2229-2236.
33. Singh, S. K., Afonina, N., Awwad, M., Bechtold-Peters, K., Blue, J. T., Chou, D., Cromwell, M., Krause, H. J., Mahler, H. C., Meyer, B. K., Narhi, L., Nesta, D. P. & Spitznagel, T. (2010). An industry perspective on the monitoring of subvisible particles as a quality attribute for protein therapeutics. *J. Pharm. Sci.* **99**, 3302-3321.
34. Ewert, S., Honegger, A. & Plückthun, A. (2003). Structure-based improvement of the biophysical properties of immunoglobulin VH domains with a generalizable approach. *Biochemistry* **42**, 1517-1528.
35. Honegger, A. & Plückthun, A. (2001). Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling and analysis tool. *J. Mol. Biol.* **309**, 657-670.
36. Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wölle, J., Plückthun, A. & Virnekäs, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J. Mol. Biol.* **296**, 57-86.
37. Anelli, T. & Sitia, R. (2008). Protein quality control in the early secretory pathway. *EMBO J.* **27**, 315-327.
38. Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D. & Duhr, S. (2010). Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* **1**, 100.
39. Schaefer, J. V. & Plückthun, A. (2012). Engineering Aggregation Resistance in IgG by Two Independent Mechanisms: Lessons from Comparison of Pichia pastoris and Mammalian Cell Expression. *J. Mol. Biol.* **417**, 309–335.
40. Liu, H., Gaza-Bulseco, G., Chumsae, C. & Newby-Kew, A. (2007). Characterization of lower molecular weight artifact bands of recombinant monoclonal IgG1 antibodies on non-reducing SDS-PAGE. *Biotechnology Letters* **29**, 1611-1622.
41. Chen, X., Tang, K., Lee, M. & Flynn, G. C. (2008). Microchip assays for screening monoclonal antibody product quality. *Electrophoresis* **29**, 4993-5002.
42. Wörn, A. & Plückthun, A. (2001). Stability engineering of antibody single-chain Fv fragments. *J. Mol. Biol.* **305**, 989-1010.
43. Ewert, S., Cambillau, C., Conrath, K. & Plückthun, A. (2002). Biophysical properties of camelid V_{HH} domains compared to those of human V_H3 domains. *Biochemistry* **41**, 3628-3636.
44. Monsellier, E. & Bedouelle, H. (2006). Improving the stability of an antibody variable fragment by a combination of knowledge-based approaches: validation and mechanisms. *J. Mol. Biol.* **362**, 580-593.
45. Buchner, J., Renner, M., Lilie, H., Hinz, H. J., Jaenicke, R., Kiefhaber, T. & Rudolph, R. (1991). Alternatively folded states of an immunoglobulin. *Biochemistry* **30**, 6922-6929.
46. Vermeer, A. W. & Norde, W. (2000). The thermal stability of immunoglobulin: unfolding and aggregation of a multi-domain protein. *Biophys. J.* **78**, 394-404.
47. Feige, M. J., Hendershot, L. M. & Buchner, J. (2010). How antibodies fold. *Trends Biochem. Sci.* **35**, 189-198.
48. Demarest, S. J., Chen, G., Kimmel, B. E., Gustafson, D., Wu, J., Salbato, J., Poland, J., Elia, M., Tan, X., Wong, K., Short, J. & Hansen, G. (2006). Engineering stability into Escherichia coli secreted Fabs leads to increased functional expression. *Protein Eng. Des. Sel.* **19**, 325-336.
49. Knappik, A. & Plückthun, A. (1995). Engineered turns of a recombinant antibody improve its *in vivo* folding. *Protein Eng.* **8**, 81-89.
50. Nieba, L., Honegger, A., Krebber, C. & Plückthun, A. (1997). Disrupting the hydrophobic patches at the antibody variable/constant domain interface: Improved *in vivo* folding and physical characterization of an engineered scFv fragment. *Protein Eng.* **10**, 435-444.
51. Niesen, F. H., Berglund, H. & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212-2221.
52. Bruylants, G., Wouters, J. & Michaux, C. (2005). Differential scanning calorimetry in life science: thermodynamics, stability, molecular recognition and application in drug design. *Curr. Med. Chem.* **12**, 2011-2020.
53. Kilar, F. & Zavodszky, P. (1987). Non-covalent interactions between Fab and Fc regions in immunoglobulin G molecules. Hydrogen-deuterium exchange studies. *Eur. J. Biochem.* **162**, 57-61.
54. Lilie, H. (1997). Folding of the Fab fragment within the intact antibody. *FEBS Lett.* **417**, 239-242.

55. Ionescu, R. M., Vlasak, J., Price, C. & Kirchmeier, M. (2008). Contribution of variable domains to the stability of humanized IgG1 monoclonal antibodies. *J. Pharm. Sci.* **97**, 1414-1426.
56. Monera, O. D., Kay, C. M. & Hodges, R. S. (1994). Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. *Protein Sci.* **3**, 1984-1991.
57. Taylor, F. R., Prentice, H. L., Garber, E. A., Fajardo, H. A., Vasilyeva, E. & Blake Pepinsky, R. (2006). Suppression of sodium dodecyl sulfate-polyacrylamide gel electrophoresis sample preparation artifacts for analysis of IgG4 half-antibody. *Anal. Biochem.* **353**, 204-208.
58. Wypych, J., Li, M., Guo, A., Zhang, Z., Martinez, T., Allen, M. J., Fodor, S., Kelner, D. N., Flynn, G. C., Liu, Y. D., Bondarenko, P. V., Ricci, M. S., Dillon, T. M. & Balland, A. (2008). Human IgG2 antibodies display disulfide-mediated structural isoforms. *J. Biol. Chem.* **283**, 16194-16205.
59. Hansen, R. E., Ostergaard, H., Norgaard, P. & Winther, J. R. (2007). Quantification of protein thiols and dithiols in the picomolar range using sodium borohydride and 4,4'-dithiodipyridine. *Anal. Biochem.* **363**, 77-82.
60. Zhang, W. & Czupryn, M. J. (2002). Free sulfhydryl in recombinant monoclonal antibodies. *Biotechnol. Prog.* **18**, 509-513.
61. Wang, Y., Lu, Q., Wu, S. L., Karger, B. L. & Hancock, W. S. (2011). Characterization and comparison of disulfide linkages and scrambling patterns in therapeutic monoclonal antibodies: using LC-MS with electron transfer dissociation. *Anal. Chem.* **83**, 3133-3140.
62. Alexander, A. J. & Hughes, D. E. (1995). Monitoring of IgG antibody thermal stability by micellar electrokinetic capillary chromatography and matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* **67**, 3626-3632.
63. Brody, T. (1997). Multistep denaturation and hierarchy of disulfide bond cleavage of a monoclonal antibody. *Anal. Biochem.* **247**, 247-256.
64. Lesk, A. M. & Chothia, C. (1988). Elbow motion in the immunoglobulins involves a molecular ball-and-socket joint. *Nature* **335**, 188-190.
65. Pritsch, O., Hudry-Clergeon, G., Buckle, M., Petillot, Y., Bouvet, J. P., Gagnon, J. & Dighiero, G. (1996). Can immunoglobulin C(H)1 constant region domain modulate antigen binding affinity of antibodies? *J. Clin. Invest.* **98**, 2235-2243.
66. Morrison, S. L., Porter, S. B., Trinh, K. R., Wims, L. A., Denham, J. & Oi, V. T. (1998). Variable region domain exchange influences the functional properties of IgG. *J. Immunol.* **160**, 2802-2808.
67. Chan, L. A., Phillips, M. L., Wims, L. A., Trinh, K. R., Denham, J. & Morrison, S. L. (2004). Variable region domain exchange in human IgGs promotes antibody complex formation with accompanying structural changes and altered effector functions. *Mol. Immunol.* **41**, 527-538.
68. Simon, T. & Rajewsky, K. (1990). Antibody domain mutants demonstrate autonomy of the antigen binding site. *EMBO J.* **9**, 1051-1056.
69. Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*. 3rd edit, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
70. Cregg, J. M., Tolstorukov, I., Kusari, A., Sunga, J., Madden, K. & Chappell, T. (2009). Expression in the yeast Pichia pastoris. *Methods Enzymol.* **463**, 169-189.
71. Bass, S., Gu, Q. & Christen, A. (1996). Multicopy suppressors of prc mutant Escherichia coli include two HtrA (DegP) protease homologs (HhoAB), DksA, and a truncated R1pA. *J. Bacteriol.* **178**, 1154-1161.
72. Cesaro-Tadic, S., Lagos, D., Honegger, A., Rickard, J. H., Partridge, L. J., Blackburn, G. M. & Plückthun, A. (2003). Turnover-based in vitro selection and evolution of biocatalysts from a fully synthetic antibody library. *Nat. Biotechnol.* **21**, 679-685.
73. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
74. Dillon, T. M., Bondarenko, P. V., Rehder, D. S., Pipes, G. D., Kleemann, G. R. & Ricci, M. S. (2006). Optimization of a reversed-phase high-performance liquid chromatography/mass spectrometry method for characterizing recombinant antibody heterogeneity and stability. *J. Chromatogr. A* **1120**, 112-120.
75. Volkova, K. D., Kovalska, V. B. & Yarmoluk, S. M. (2007). Modern techniques for protein detection on polyacrylamide gels: problems arising from the use of dyes of undisclosed structures for scientific purposes. *Biotech. Histochem.* **82**, 201-208.

Supplementary Material for Chapter 3.1

Construction of Expression Plasmids

Expression vectors for mammalian HEK293 cells were first designed for the independent expression of heavy or light chain, respectively, using the pMORPH® vector series (MorphoSys).[1] The V_H and V_L genes of the previously analyzed scFv fragments[2] were amplified and inserted into pMORPH®2_h_IgG1f, pMORPH®2_h_Igκ or the pMORPH®2_h_Igλ2 vectors. These vectors led to the expression of correctly assembled full-length IgG into the supernatant of mammalian cells upon co-transfection. These proteins carry no tags. The creation of Fab expression vectors was performed likewise, in this case however, the V_H/C_H1 chain was fused to a myc-tag and $(his)_6$-tag at its C-terminus for detection and purification purposes.

For the establishment of stable cell lines using the Flp-In system (Invitrogen), the two expression cassettes for H and L, respectively, were combined in a single vector based on the pcDNA5/FRT series (Invitrogen). After removal of the intrinsic *Mfe*I cleavage site by site-directed mutagenesis (using primers MfeI_del_for: 5′ - GGC AAG GCT TGA CCG ACA GTT GCA TGA AGA ATC TGC - 3′ and MfeI_del_rev: 5′ - GCA GAT TCT TCA TGC AAC TGT CGG TCA AGC CTT GCC - 3′ to create the pcDNA5/FRT/-MfeI plasmid), the final vectors were constructed as follows. The heavy chain gene was amplified from the corresponding pMORPH® vector using the primers HC_NheI_for (5′ - CCC AAG CTG *GCT AGC* GCC ACC ATG AAA CAC C - 3′) and HC_XbaI_rev (5′ - GCG TCC *TCT AGA* GCG TAC CCA ATT CAA CAG GC - 3′). The resulting amplicon contained the complete coding sequence of the heavy chain as well as a poly A sequence derived from the human growth hormone, but no CMV promoter. The amplification of the light chain gene was performed using LC_XbaI_for (5′ - GCG TCC *TCT AGA* CGC GAT GTA CGG GCC AGA TAT ACG - 3′) and LC_ApaI_rev (5′ - GCC ACC CGT TTA AAC *GGG CCC* C - 3′), resulting in an amplicon containing a CMV promoter, the complete coding sequence of the light chain, but no poly A sequence. Regions complementary to the sequences found in the parental vectors are underlined, while nucleotides allowing the cleavage by *Nhe*I, *Xba*I or *Apa*I are printed in italics. After double-digesting the heavy chain amplicon with *Nhe*I-*Xba*I and the light chain amplicon with *Xba*I-*Apa*I, the resulting fragments were ligated into the *Nhe*I-*Apa*I digested pcDNA5/FRT/-MfeI plasmid. Since this vector already contained one intrinsic CMV promoter and a poly A sequence (derived from bovine growth hormone), both the sequences coding for the heavy as well as for the light chain were finally present as complete expression cassettes including CMV promoter and poly A sequences of their own. The final vectors were sequenced and tested by transient transfection for their functionality before being used for the establishment of stable cell lines.

The *Pichia* expression vectors are based on the vector pGAPZαB (Invitrogen). The sequences for the light chains and the different heavy chain variants were amplified from the pMORPH® vectors described above. Since for expression in *Pichia pastoris* a plasmid has to be integrated into the genome, also these genes were combined within one single vector, retaining them as independent

expression cassettes, each with its own GAP promoter and termination sequence. To assemble the final vectors, the cloning of intermediate vectors was necessary. First, two pGAPZαB derivatives were constructed for each construct: one containing the gene encoding the corresponding light chain, the other one holding the heavy chain gene as complete expression cassettes. All expression vectors were obtained by inserting a PCR-amplified fragment from the pMORPH® vectors containing the sequence coding for the mature heavy or light chain into the multiple cloning site (MCS) of pGAPZαB. In these vectors, the POI is placed behind the α-factor pre-pro (αMFpp) sequence and under control of the constitutive GAP promoter.

As restriction enzymes, *EcoR*I introduced downstream of the N-termini and *Xba*I introduced upstream of the C-termini of the corresponding genes were used. The heavy chain fragment was amplified using either PP_WT_for (5′ - GCG TCC *GAA TTC* AA CAG GTG CAA TTG CAA CAG TC - 3′) or PP_M_for (5′ - GCG TCC *GAA TTC* AA CAG GTG CAA TTG GTA CAG TC - 3′) in combination with PP_HC_rev (5′ - GCG TCC *TCT AGA* TCA CTA TCA TTT ACC CGG AGA CAG GG - 3′). The light chain fragments were created in the analogous manner using the oligonucleotides PP_λ_for (5′ - GCG TCC *GAA TTC* AA GAT ATC GAA CTG ACC CAG CC - 3′) and PP_λ_rev (5′ - GCG TCC *TCT AGA* TCA CTA TGA ACA TTC TGT AGG GGC CAC - 3′) for the lambda light chain and PP_κ_for (5′ - GCG TCC *GAATTC* AA GAT ATC GTG CTG ACC CAG AG - 3′) and PP_κ_rev (5′ - GCG TCC *TCTAGA* TCA CTA AAC ACT CTC CCC TGT TGA AGC TC - 3′) for the kappa counterpart. Regions complementary to the N- or C-termini of the heavy or light chain, respectively, are underlined, while nucleotides allowing the cleavage by *EcoR*I or *Xba*I are printed in italics. Two consecutive additional stop codons were introduced into the construct by the dotted underlined triplets. The AA doublet in the forward primers completed to original sequence from the MCS to encode the four additional residues AGIQ upstream of the mature N-terminus of either the light or heavy chains, respectively. These additional amino acids were removed in a later step from the heavy chain construct by assembly PCR using the following primers: WT_AGIQdel_for (5′ - GA GAG GCT GAA GCT CAG GTG CAA TTG CAA CAG TCT GGT CCG GG - 3′) and WT_AGIQdel_rev (5′ - G CAA TTG CAC CTG AGC TTC AGC CTC TCT TTT CTC GAG AGA TAC CC - 3′) or M_AGIQdel_for (5′ - GA GAG GCT GAA GCT CAG GTG CAA TTG GTA CAG TCT GGT CCG GG - 3′) and M_AGIQdel_rev (5′ - C CAA TTG CAC CTG AGC TTC AGC CTC TCT TTT CTC GAG AGA TAC CC - 3′). The AGIQ-coding sequences of the light chains were removed by conventional cloning using the unique *Xho*I-*Nco*I cleavage sites, employing the oligonucleotides λ_AGIQdel_for (5′ - GCG TCG *CTC GAG* AAA AGA GAG GCT GAA GCT GAT ATC GAA CTG ACC CAG CC - 3′) or κ_AGIQdel_for (5′ - GCG TCG *CTC GAG* AAA AGA GAG GCT GAA GCT GAT ATC GTG CTG ACC CAG AGC - 3′) in combination with LC_AGIQdel_rev (5′ - CCG GAA CGG CAC TGG TCA ACT TGG C - 3′). Regions complementary to the N-termini of the light chains are underlined, while nucleotides allowing the cleavage by *Xho*I or *Nco*I are printed in italics.

Finally, the two vectors (each coding for either the heavy or light chain, respectively) were combined in a single vector for integration into the *Pichia* genome. This step was performed on the basis of the methods described by Ogunjimi and co-workers.[3] Briefly, the vectors containing the heavy or light chain gene was double-digested with either *Mlu*I-*BamH*I or *Mlu*I-*Bgl*II, respectively. As digestion with *BamH*I-*Bgl*II resulted in compatible overhangs, the subsequent ligation eliminated the used *Bgl*II

cleavage site, leading to only one remaining *Bgl*II cleavage site in the final plasmid, which was subsequently used for linearization of the plasmid for genomic integration. This vector contained both independent expression cassettes but only one origin of replication as well as only one Zeocin™ resistance marker.

1. Steidl, S., Ratsch, O., Brocks, B., Durr, M. & Thomassen-Wolf, E. (2008). In vitro affinity maturation of human GM-CSF antibodies by targeted CDR-diversification. *Mol. Immunol.* **46**, 135-144.
2. Ewert, S., Honegger, A. & Plückthun, A. (2003). Structure-based improvement of the biophysical properties of immunoglobulin VH domains with a generalizable approach. *Biochemistry* **42**, 1517-1528.
3. Ogunjimi, A. A., Chandler, J. M., Gooding, C. M., Recinos, I. A. & Choudary, P. V. (1999). High-level secretory expression of immunologically active intact antibody from the yeast Pichia pastoris. *Biotechnology Letters* **21**, 561-567.

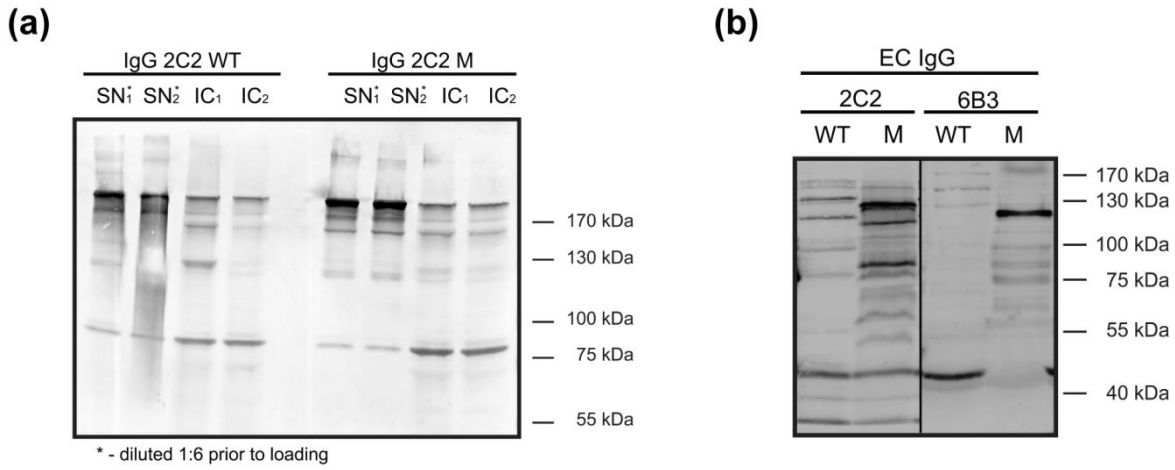


Fig. S3.1.1. Comparison of secreted IgG amounts with intracellular levels and IgG expression in *E. coli*. (a) Western blot analysis of IgG 2C2 WT and M variants found in the supernatant (SN) and intracellularly (IC) in stable HEK cells. Samples were taken in duplicates and the cells were lysed with RIPA-lysis buffer containing 1% NP40 and Na-deoxycholate. The supernatant was diluted 6x to allow the detection of the significantly lower intracellular IgG levels on the same blot. Samples were heated to 95°C for 5 min and the detection performed with antibodies specific for the respective light chain. **(b)** Western blot analysis of full-length IgGs expressed in *E. coli*. The OD-normalized periplasm was isolated and separated by non-reducing SDS-PAGE.

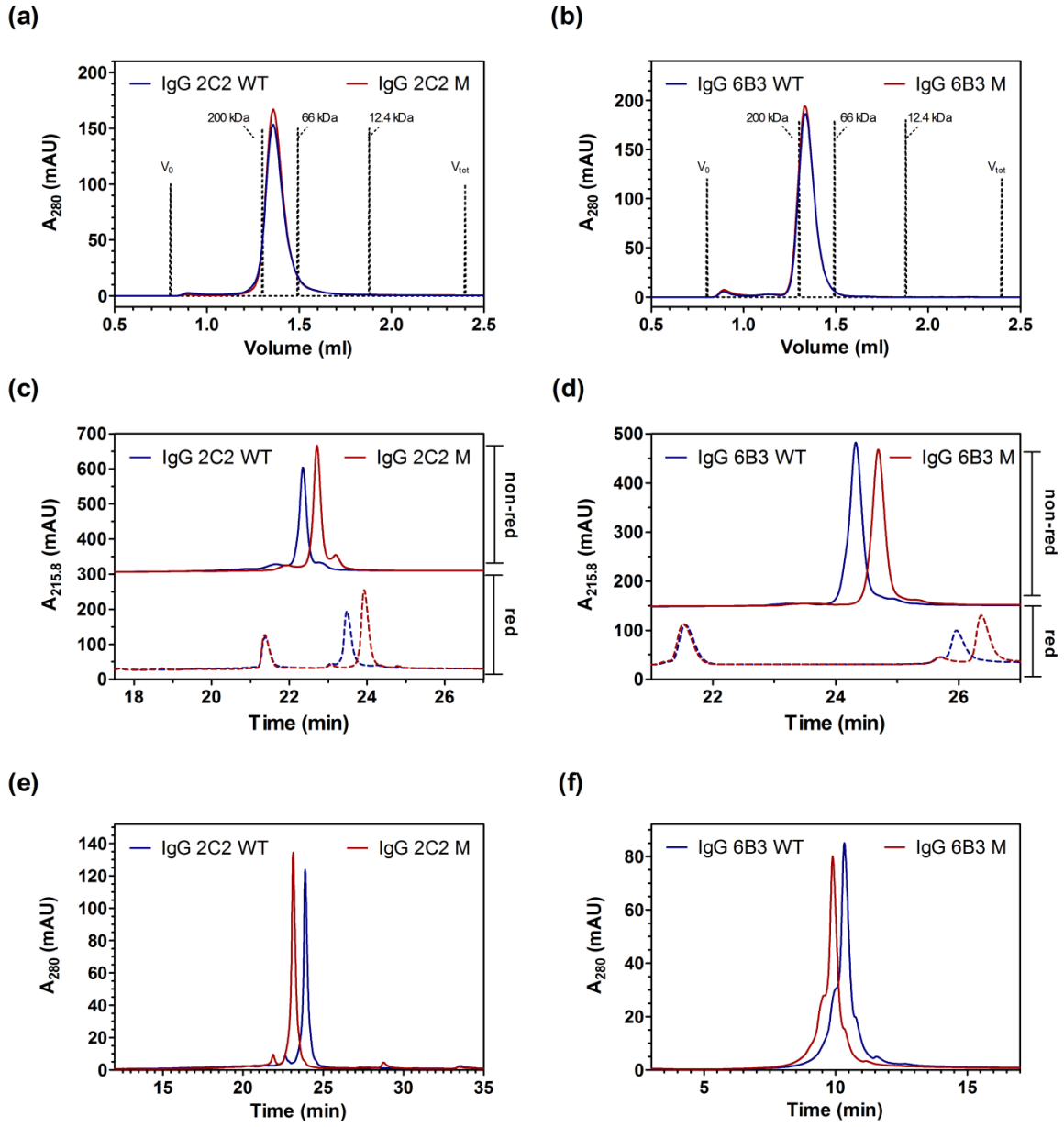


Fig. S3.1.2. Chromatographic analysis of HEK produced IgG variants. Signals derived from WT IgG are shown in blue, those of the mutant M in red. **(a)** Size exclusion chromatography (SEC) analysis of 2C2 variants. Dotted lines indicate elution volumes of molecular mass standards (from left to right): β-amylase (200 kDa), bovine serum albumin (66 kDa) and cytochrome c (12.4 kDa). Also shown are the void volume V_0 and the total volume V_{tot}. **(b)** SEC of IgG 6B3 variants analyzed analogous to panel (a). **(c)** Analysis of IgG 2C2 WT and M by reverse phase HPLC (RP-HPLC). Changes in the elution profiles are attributed to the increased hydrophobicity of the M variant. The upper panel shows the chromatogram recorded under non-reducing conditions, the lower one after reduction. **(d)** RP-HPLC of IgG 6B3 variants analog to panel (c). **(e)** Cation exchange chromatography (CIEX) of IgG 2C2 variants. Differences in the elution profile of the M variant are caused by the additional negative charge introduced by the V72D mutation. **(f)** CIEX of IgG 6B3 variants analog to panel (e).

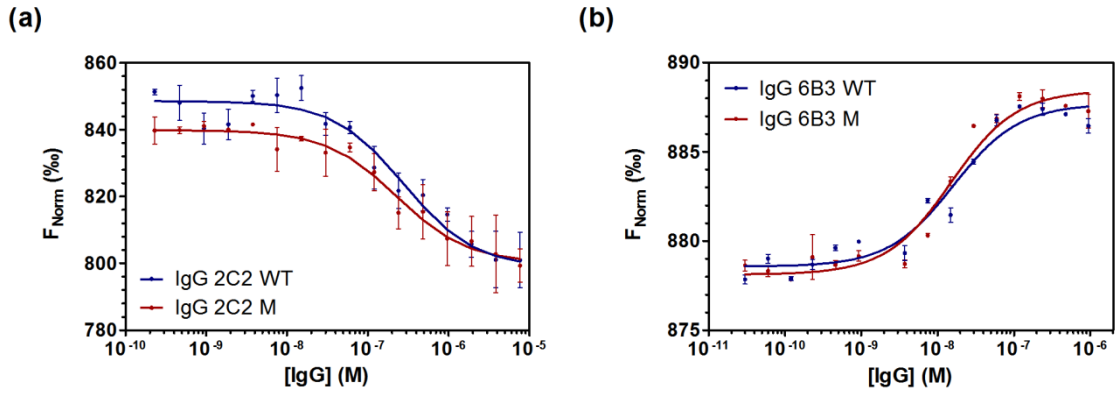


Fig. S3.1.3. Binding affinity of IgG mutants obtained from mammalian expression. Direct interaction studies between purified IgGs and their fluorescently labeled antigens (M18-transferrin for IgG 2C2 and myoglobin for IgG 6B3, respectively) analyzed using microscale thermophoresis technology. Normalized fluorescence [‰] is plotted against the IgG concentration titrated from 0.03 nM to 10,000 nM. The fitted K_D values lie within the experimental error range and indicate that the binding is fully retained for **(a)** the IgG 2C2 variants and **(b)** IgG 6B3 WT and M, respectively.

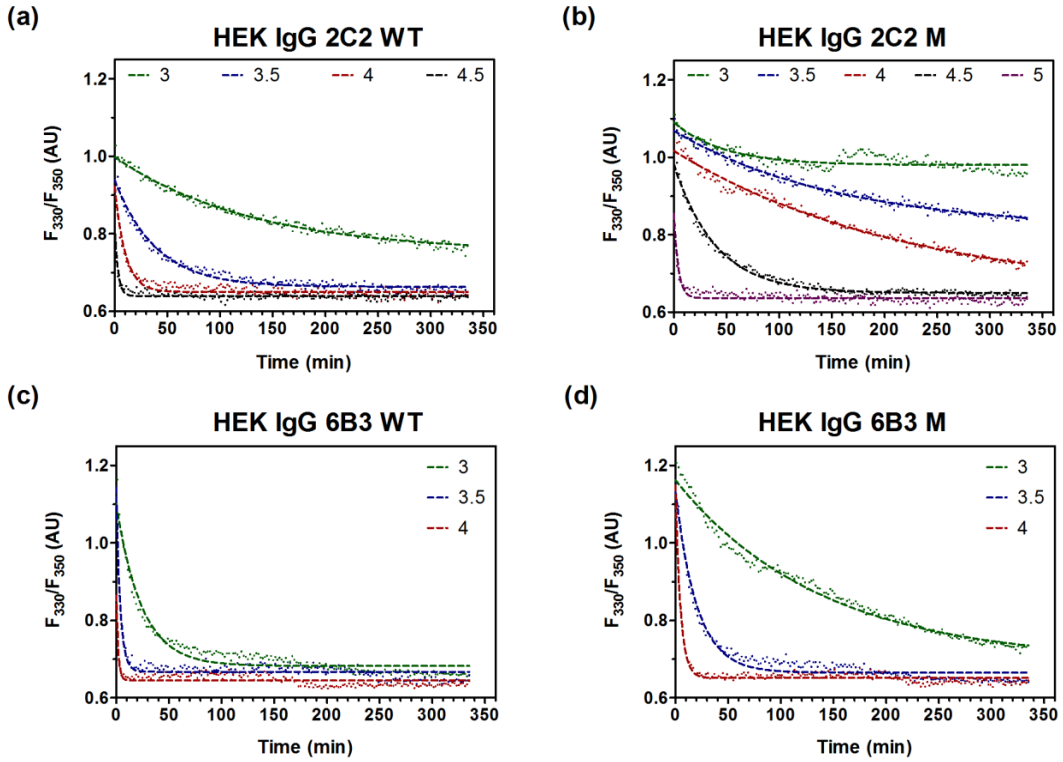


Fig. S3.1.4. Real-time analysis of GdnHCl-induced denaturation of IgGs. The graphs were obtained from the F_{330}/F_{350} ratio plotted as a function of time. Shown are the signals derived from IgG sample in the presence of various GdnHCl concentrations. Real-time unfolding was recorded over a period of 5.5 h for **(a)** HEK IgG 2C2 WT, **(b)** HEK IgG 2C2 M, **(c)** HEK IgG 6B3 WT and **(d)** HEK IgG 6B3 M.

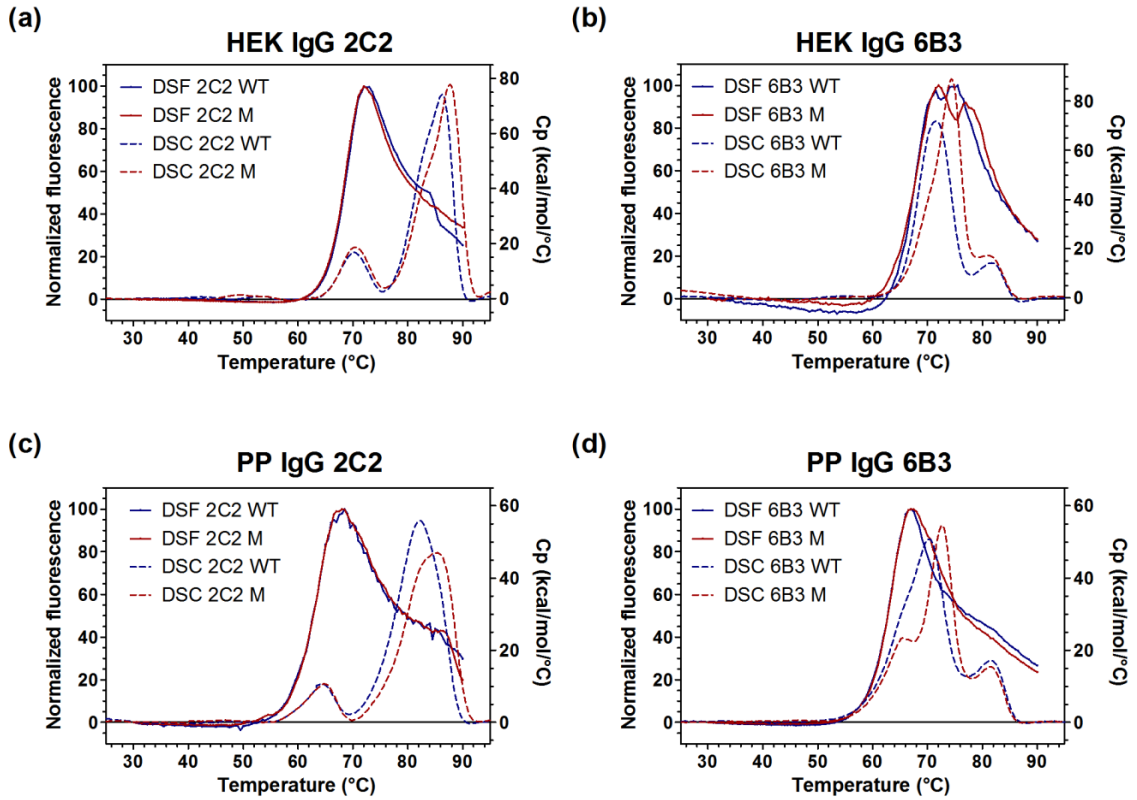


Fig. S3.1.5. Comparison of DSF and DSC data from figures 3.1.6 and 3.1.7. To allow an eased comparison of domain stabilities, data recorded by DSF (displayed as continuous lines) and DSC (dotted lines) are combined within one single graph. Comparative graphs were generated for **(a)** HEK IgG 2C2, **(b)** HEK IgG 6B3, **(c)** PP IgG 2C2 and **(d)** PP IgG 6B3.

			ITF	GdnHCl	DSF	DSC
IgG 2C2	**WT**		70.4°C*	2.5 M	n.d.	86.0°C
	M		71.8°C*	3.8 M	n.d.	87.8°C
		Δ =	**1.4°C**	**1.3 M**	**-**	**1.8°C**
IgG 6B3	**WT**		67.6°C	2.0 M	74.5°C	72.1°C
	M		70.8°C	2.6 M	77.0°C	74.3°C
		Δ =	**3.2°C**	**0.6 M**	**2.5°C**	**2.2°C**
Fab 6B3	**WT**		69.7°C	2.0 M	76.5°C	72.6°C
	M		74.2°C	2.6 M	80.0°C	76.6°C
		Δ =	**4.5.°C**	**0.6 M**	**3.5°C**	**4.0°C**

Fig. S3.1.6. Overview of stabilities determined by different methods. Indicated are the melting temperatures measured in Centigrade (°C) as well as the difference between the M and WT construct (Δ). For the DSF and DSC data, the temperatures derived from the respective Fab fragment are listed. For the denaturant-induced unfolding, the GdnHCl concentration (in M) is specified which correlates to the midpoint of transition of the corresponding graphs shown in Fig. 3.1.5.

3.2 Engineering aggregation resistance in IgG by two independent mechanisms: lessons from comparison of *Pichia pastoris* and mammalian cell expression

Jonas V. Schaefer and Andreas Plückthun

in **Journal of Molecular Biology** (2012) 417, 309–335

Abstract

Aggregation is an important concern for therapeutic antibodies, since it can lead to reduced bioactivity and increase the risk of immunogenicity. In our analysis of immunoglobulin G molecules (IgGs) of identical amino acid sequence but produced either in mammalian cells (HEK293) or the yeast *Pichia pastoris*, dramatic differences in their aggregation susceptibilities were encountered. The antibodies produced in *Pichia* were much more resistant to aggregation under many conditions, a phenomenon found to be mainly caused by two factors. First, the mannose-rich glycan of the IgG from *Pichia*, while slightly thermally destabilizing the IgG, strongly inhibited its aggregation susceptibility, compared to the complex mammalian glycan. Second, on the *Pichia* produced IgGs amino acids belonging to the α-factor pre-pro sequence were left at the N-termini of both chains. These additional residues proved to considerably increase the temperature of the onset of aggregation and reduced the aggregate formation after extended incubation at elevated temperatures. The attachment of these residues to IgGs produced in cell culture confirmed their beneficial effect on the aggregation resistance. Secretion of IgGs with native N-termini in the yeast system became possible after systematic engineering of the precursor proteins and the processing site. Taken together, the present results will be useful for the successful production of full-length IgGs in *Pichia pastoris*, give indications on how to engineer aggregation-resistant IgGs and shed new light on potential biophysical effects of tag sequences in general.

Keywords: IgG, aggregation susceptibility, antibody engineering, *Pichia pastoris*, GAP promoter, signal sequence, α-factor pre-pro sequence

Abbreviations used:

αMFpp, α-factor pre-pro sequence; CD, Circular dichroism; CMV, Cytomegalovirus; DMEM, Dulbecco's modified Eagle medium; DMSO, Dimethyl sulfoxide; DSC, Differential scanning calorimetry; ELISA, Enzyme-linked immunosorbent assay; FcγR, Fc gamma receptors; FBS, Fetal bovine serum; FDA, Food and Drug Administration; GdnHCl, Guanidine hydrochloride; H, Heavy chain of an antibody; HEK cells, Human embryonic kidney cells; HT, High-tension voltage; IgG, Immunoglobulin G; L, Light chain of an antibody; MALS, Static multi-angle light scattering; MRE, Mean residue ellipticity; MS, Mass spectrometry; MST, Microscale thermophoresis; PBS, Phosphate buffered saline; pI, Isoelectric point; POI, Protein of interest; PP, *Pichia pastoris;* scFv, Single-chain variable fragment of an antibody; SEC, Size-exclusion chromatography; SDS-PAGE, Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

Introduction

Antibodies and their derivatives have found a broad range of applications, from basic research to medical therapy. Using many different techniques for their generation[1] several therapeutic immunoglobulins, usually in the whole IgG format, have been FDA-approved and applied to the treatment of a wide range of diseases.[2-4] One major challenge encountered in the development of antibody-based therapeutics, however, is their aggregation susceptibility under both formulation conditions[5; 6] and in vivo, where aggregation is involved in a significant degradation pathway for monoclonal antibodies.[7] Irreversible aggregation thereby leads to reduced bioactivity[8] and presents an increased risk of immunogenicity.[9-14] Aggregation of even only a few percent of the administered therapeutic antibody can cause severe problems, as already low amounts of aggregates have long been known to cause anaphylactoid reactions.[15; 16] Therefore, engineering IgG to reduce or eliminate its self-association is of great importance and might further extend the potential of therapeutic antibodies.

Since nearly all recombinant therapeutic antibodies are expressed in the IgG format, choices must be made in the molecular composition and expression system. Currently, all approved therapeutic antibodies are produced in mammalian cells,[17] since only this expression system can introduce the complex glycosylation which is necessary for the interaction of the antibodies with some of the activating Fcγ receptors.[18] IgGs normally possess a single N-linked glycosylation at Asn297 in the C_H2 domain, maintaining these domains at a critical distance to bind to the Fcγ receptors.

In some cases, however, this FcγR interaction is neither needed nor wanted, thus opening the range of possible production hosts also to *Pichia pastoris*. In the past, this methylotrophic yeast has been widely used for the expression of both secreted and intracellular eukaryotic proteins.[19] While it is able to carry out efficient disulfide bond formation and isomerization using the eukaryotic secretion quality control machinery,[19-21] *Pichia* expression generates mannose-rich glycosylation of IgG.[22] The attached glycan moieties differ markedly from the product of mammalian cells, resulting in antibodies that lack FcγR binding capacity. However, recently, *Pichia pastoris* has also been engineered to introduce complex, human-like glycosylation.[23-26]

Due to its respiratory growth, *P. pastoris* can grow to very high cell densities,[27] and therefore express high levels of recombinant protein through efficient and tightly regulated promoters, accounting for up to 30% of the total cell protein.[28] Moreover, comparatively few endogenous proteins are secreted by this yeast system, which offers further advantages for subsequent protein purification and downstream processing.[29] Since Ridder and colleagues reported the first successful production of functional antibody fragments in *Pichia*,[30] more than 50 reports have been published describing antibody expression in this system.[31] While most of these molecules were scFv fragments and fusions thereof, several Fab fragments have been produced in *Pichia* as well.[32-34] Interestingly, until today *P. pastoris* has only been used as an expression host for a handful of full-length IgGs.[25; 26; 35]

Almost all heterologous proteins produced in *P. pastoris* have been expressed under the control of the methanol-inducible alcohol oxidase 1 (AOX1) promoter.[20; 36] However, throughout this study the constitutive glyceraldehyde-3-phosphate dehydrogenase (GAP) promoter[37] was used, since the ability to compare mutant proteins for their expression and aggregation tendencies was crucial. The GAP promoter is usually considered to result in lower protein yields compared to AOX1,[38] even though some reports describe the GAP promoter to be superior to AOX1 for certain proteins,[39; 40] including Fab fragments.[41] A potential reason could be the lower rate of expression which might better match the corresponding folding rate and, thus, reduces the stress level for the cells.

Previously, our laboratory analyzed sequence modifications in the variable domain in various antibody formats for aggregation propensity and conformational stability.[42-44] To carry out such studies for whole IgGs, eukaryotic expression systems are required where functional expression yields can be compared and mutant proteins can be isolated and biophysically characterized. Therefore, it is vital that isogenic strains can be produced, only differing by the sequence of the protein to be tested. As a first step, we investigated homologous recombination in HEK293 cells[45] and *Pichia pastoris*. As *P. pastoris* does not maintain episomal vectors, the expression cassettes are generally designed for integration into the yeast genome, thus leading to stable cell lines expressing the protein of interest (POI). In initial experiments, we were very surprised to find large differences between the aggregation behaviors of purified IgG protein from both systems, even though their amino acid sequence was identical: the same recombinant IgG expressed in HEK293 cells was much more aggregation-prone than its counterpart from *Pichia pastoris*.

Here, we describe our investigations of this phenomenon and pinpoint it to differences in the processing of the precursor proteins, as well as to differences between the unequal glycosylations. As an initial model system, an antibody was used that had originally been selected from the HuCAL library[46] by panning against myoglobin from horse skeletal muscle and subsequently been stability-engineered.[47] For this study, it was converted to a full-length IgG and investigated for its processing and expression in mammalian HEK293 cells and in the yeast *Pichia pastoris*.

Results

Comparison of various IgG constructs

In this study, we compared IgGs of identical amino acid sequence, but produced in different eukaryotic expression systems. Unexpectedly, these molecules revealed dramatic differences in some of their biophysical properties. To monitor changes in their structure as a function of temperature, purified IgGs were analyzed by circular dichroism (CD). Since the signals stemming from the IgG's β-sheets and random coil essentially canceled out to zero at 208 nm, thermal denaturation of the IgG molecules was determined at this wavelength as a function of temperature. These analyses indicated that the IgGs produced in human embryonic kidney cells (HEK) and in *Pichia pastoris* (PP) not only had a slightly different temperature of unfolding onset, but, unexpectedly, also showed very different aggregation susceptibilities (Fig. 3.2.1a). The abrupt jump upwards in the respective CD vs. temperature curve indicates that the antibody expressed in mammalian cells started to aggregate at a temperature higher than ~75°C, whereas the corresponding construct produced in the *Pichia pastoris* system did not exhibit any detectable aggregation at all. This finding could also be visually observed, as the mammalian antibodies had turned into a turbid sample after the experiment, compared to a clear solution for the yeast IgG (Fig. 3.2.1b, inset: PP samples left and HEK293 samples right). To further confirm these results, an aggregation assay based on light scattering (Fig. 3.2.1b) was performed in a fluorimeter: when both excitation and emission wavelengths were set to 500 nm, the different IgG molecules showed clear differences in the intensity of scattered light upon exposure to increasing temperatures. Only the molecules produced in mammalian cells aggregated, forming complexes large enough to measurably scatter the incoming light from about ~75°C onwards. In contrast, the *Pichia*-produced IgG did not show any detectable light scattering at all.

While the protein sequences of the mature IgG H and L chains are identical, different sugar moieties are attached to the C_H2-domain of the heavy chain in yeast and in mammalian cells. Therefore, glycan knockout mutants were designed by replacing the glycosylation motif Asn297-Ser298-Thr299 in the C_H2-domain by an Asn297-Ser298-Ala299 sequence. The resulting unglycosylated T299A mutants were analyzed for their behavior in the aggregation assay in comparison to their glycosylated counterparts. As can be seen in Fig. 3.2.1c, the glycan moieties did influence the temperature of IgG aggregation onset substantially - however, they seemed to be not the only cause of the different aggregation behavior: the unglycosylated IgG produced in the yeast system did show aggregation, seen as a clear increase in light scattering at 81.3 ± 0.1°C, while their unglycosylated mammalian counterparts already aggregated at 77.3 ± 0.1°C. Glycosylated HEK IgG thus aggregates at a temperature ~3.2°C higher than the non-glycosylated HEK IgG.

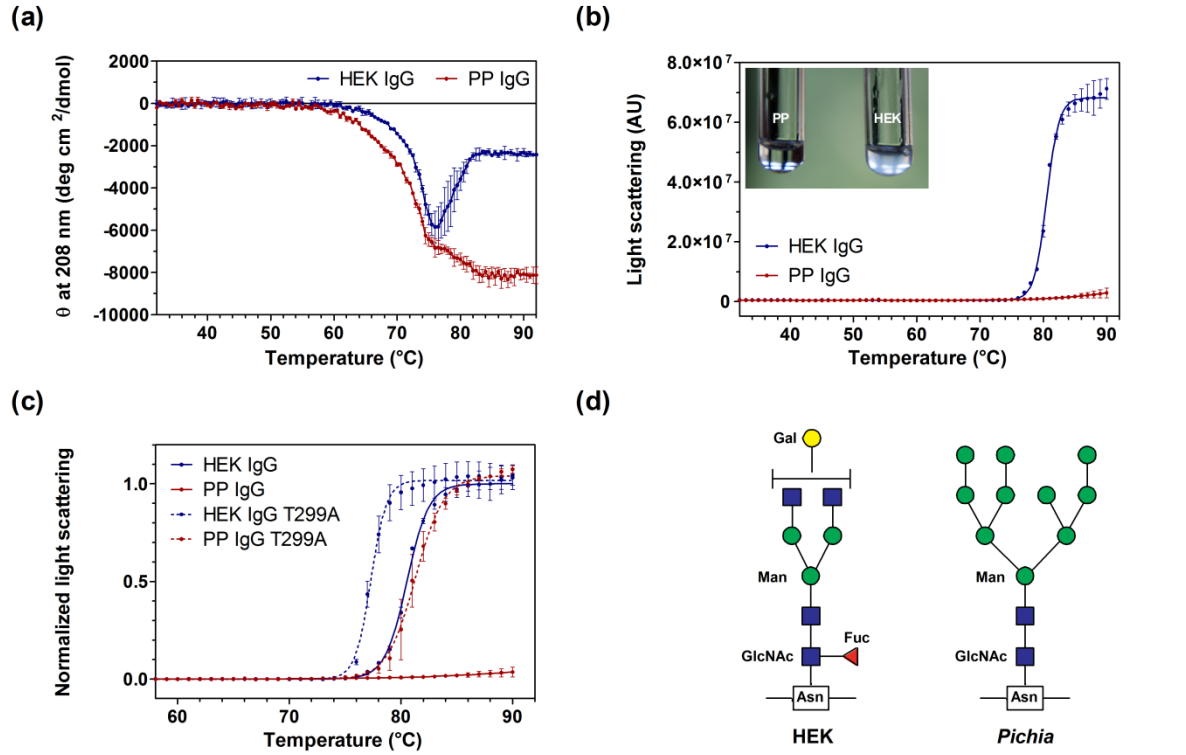


Fig. 3.2.1. Aggregation behavior of IgGs produced in eukaryotic expression systems. (a) Thermal denaturation curves of IgGs produced in mammalian cells (HEK293) or in the yeast *Pichia pastoris* (PP), respectively. The denaturation was followed by circular dichroism (CD), plotting the signals at 208 nm as a function of temperature. The values are reported as mean residue ellipticity (MRE, θ). The abrupt jump upwards at about 79°C in the HEK-derived signal is caused by the formation of insoluble aggregates. **(b)** Aggregation of IgG constructs as a function of temperature measured by light scattering at 500 nm. Emission was recorded at the same wavelength as excitation was performed. The insert shows PP (left) and HEK293 (right) samples after the measurement. **(c)** Aggregation susceptibility of glycosylated and deglycosylated (glycan knockout T299A) IgGs, analyzed by light scattering at 500 nm. **(d)** Structures of C_H2-attached glycans of IgGs produced by either the HEK or the PP system, as determined by ESI-MS. GlcNAc - N-acetyl-D-glucosamine, Man - mannose, Fuc - fucose, Gal - galactose.

We determined the different glycan structures present in the IgGs from both the mammalian and yeast expression systems by mass spectrometry (MS). This was achieved by subtracting the experimentally determined weight of the T299A constructs (49,798.5 Da) from that of the original, glycosylated variant (51,274 Da), correcting it for the Thr to Ala mutation (- 30 Da). From the resulting mass difference of 1445.5 Da (Fig. 3.2.2a), the glycan structure of HEK IgG schematically shown in Fig. 3.2.1d can be deduced - the $Gal(GlcNAc)_2(Man)_3(GlcNAc)_2Fuc$ composition typically found for proteins expressed in human cell culture.[48] The yeast glycan, on the other hand, was found to have a mass of 2,035 Da (52,399 Da - 50,334 Da – 30 Da) and to be rich in mannose units, the total number of which varied between 9 and 15, with the majority of molecules possessing 10 mannose units (molecular formula $(Man)_{9\text{-}15}(GlcNAc)_2$. These different glycan structures between *Pichia* and mammalian cells were found to be the reason for the slightly decreased temperature of the onset of unfolding recorded in Fig. 3.2.1a for the *Pichia*-produced IgG. It could be shown in various experiments that the yeast glycan destabilizes the C_H2-C_H2 domain interface compared to the mammalian counterpart and thereby decreases the thermal stability of the whole IgG (Supplementary Data Fig. S3.2.1). Yet, at the same time, it largely prevents irreversible aggregation (Fig. 3.2.1c).

Since the different glycan structures were not the only cause of the different aggregation susceptibilities of the HEK- and *Pichia*-produced IgGs, both molecules were further analyzed by MS in their native and PNGase F deglycosylated form. As listed in Fig. 3.2.2a, not even in their deglycosylated forms did they correspond to the theoretical expectation from the aa sequence. However, considering that the heavy chains of mammalian IgGs generally lack their C-terminal Lys residues (-146 Da),[49] possess pyroglutamate residues at their N-termini (leading to the loss of a water molecule (-18 Da))[50; 51] and exhibit oxidation of their methionine residues (+16 Da),[52] the underlined calculated masses, which are in excellent agreement with the measured ones, offer a plausible explanation for the posttranslational modifications of the mammalian IgG heavy chain. The light chain showed the expected mass without any major post-translational modifications.

The results for the *Pichia*-produced IgGs, however, indicated the existence of large additional modifications at both the heavy and the light chain. Edman sequencing of their respective amino-termini verified the presence of the tetrapeptide Glu-Ala-Glu-Ala (EAEA) in front of the genuine N-terminus of the proteins, accounting for the additional 400 Da measured by MS (double underlined). Minor fractions with one or both EA-pairs deleted were not detected. Since these four amino acids are remains of the α-factor pre-pro sequence (αMFpp) used in the expression setup, the maturation process of secreted proteins within the yeast system was analyzed in more detail. This secretion system consists of a 19-aa signal (pre) sequence followed by a 66-residue (pro) sequence, the latter containing three consensus N-linked glycosylation sites.[53] Directly following a dibasic Kex2 endopeptidase processing site, a spacer peptide of four residues (EAEA) is located upstream of the genuine N-terminus of the mature α factor peptide. This spacer was included in the present IgG construct as well. As schematically illustrated in Fig. 3.2.2b for the heavy chain, proteins fused to the αMFpp sequence undergo various processing steps until being fully matured. Several studies have shown that dipeptidyl aminopeptidase A (DPAPase A) has problems in fulfilling its task in protein overexpression, thereby often leaving residual EAEA residues on the protein to be secreted.[54-56] This also was the case for the *Pichia*-produced IgG of this study, which therefore will be referred to "H-E / L-E" from now on (with the suffix "E" representing the EAEA tetrapeptide), having this extension on both chains.

N-terminal engineering of *Pichia* constructs

To investigate whether antibody chains with genuine N-termini can be produced in *Pichia pastoris*, yeast variants were designed lacking the EAEA tetrapeptide on the mature heavy and light chain, respectively. The EAEA coding sequence was deleted from the DNA encoding both chains - either in the individual coding sequences or for both chains simultaneously (depicted in Fig. 3.2.2d, exemplified for the light chain).

(a)

construct	measured weight *	PNGaseF treated §	theoretical weight #	theoretical - Lys $	theoretical - Lys - H_2O ß	theoretical - Lys - H_2O + Ox &	theoretical + EAEA %	Δ mass ¶
heavy chain (H)								
HEK	51,274	49,829	49,977	49,831	49,813	49,829	50,377	0
PP	52,399	50,376						1
light chain (L)								
HEK	22,521	22,521	22,525	-	-	-	22,925	4
PP	22,922	22,922						3

* - mass (Da) measured by ESI-MS under reducing conditions

§ - mass (Da) measured by ESI-MS under reducing conditions after deglycosylation by PNGase F

- mass (Da)calculated from the sequence of the unmodified, reduced mature chain

$ - mass (Da) calculated considering truncation of C-terminal Lys residue

ß - mass (Da) calculated considering truncation of C-terminal Lys residue and pyroglutamate formation (loss of a water molecule)

& - mass (Da) calculated considering truncation of C-terminal Lys residue, pyroglutamate formation and Met oxidation

% - mass (Da) calculated considering N-terminal extention by Glu-Ala-Glu-Ala (no pyroglutamate formation)

¶ - mass difference between the best matching theoretical value and the PNGase F treated measured value (underlined or double underlined)

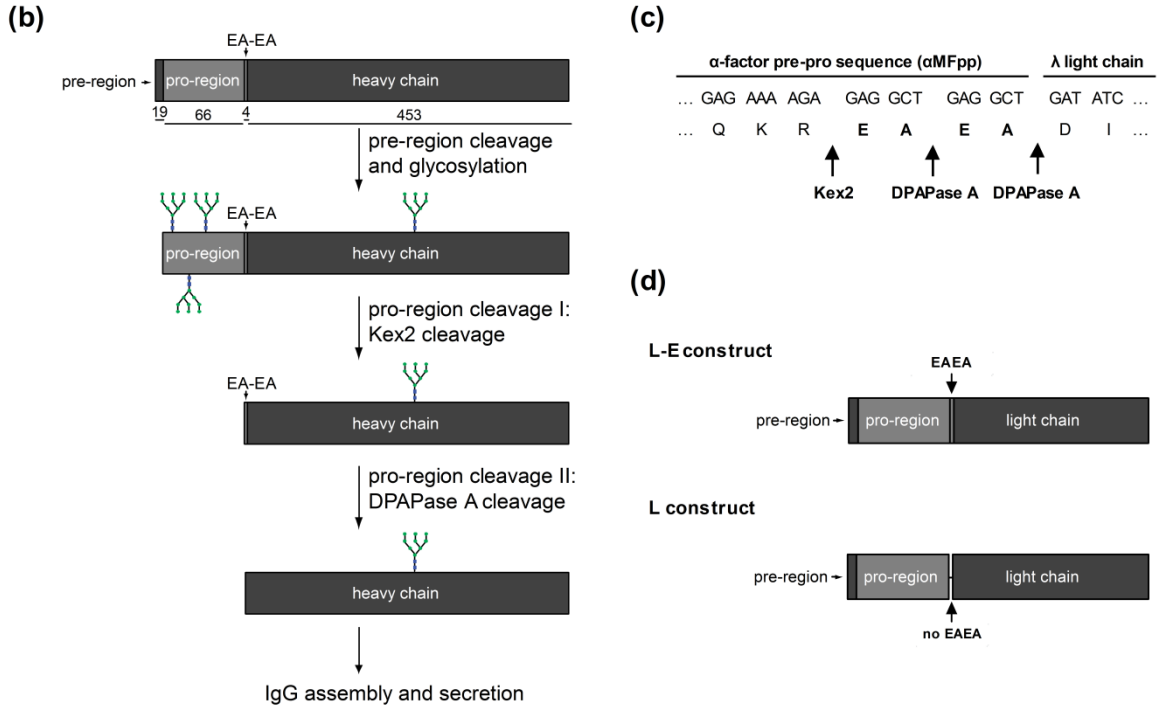
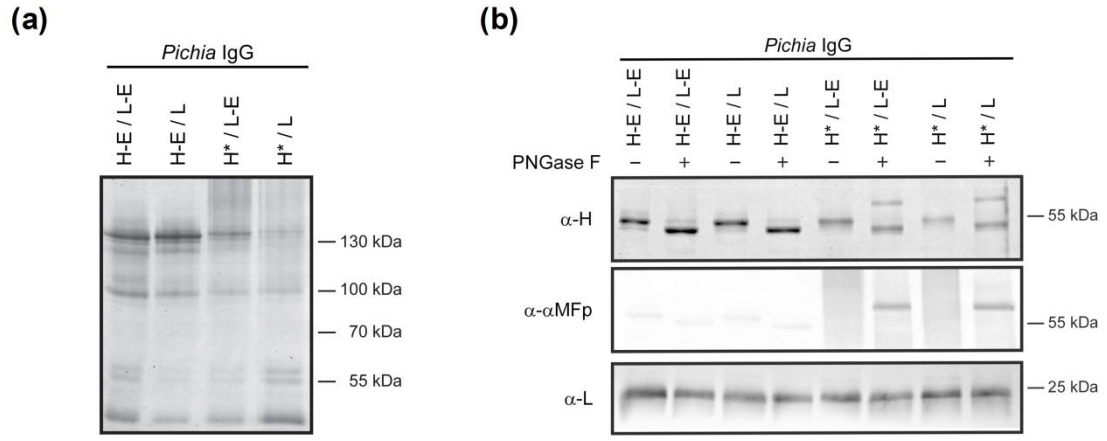


Fig. 3.2.2. Overview of detected IgG masses and *Pichia* processing. (a) Summary of molecular masses of HEK- and *Pichia*-produced IgGs, analyzed under reducing conditions by ESI-MS and calculated based on the amino acid sequence of the individual protein chains. All masses are stated in Da. **(b)** Schematic overview (not drawn to scale) of N-terminal processing of secreted proteins fused to the α-factor pre-pro sequence (αMFpp) by the endopeptidase Kex2 and the dipeptidyl aminopeptidase A (DPAPase A, product of the STE13 gene). The residue numbers of the different regions of the precursor protein are shown below the top construct. After the pre-region is cleaved off by signal peptidase upon entering the endoplasmic reticulum, the pro-region containing protein is glycosylated and is further processed in the Golgi apparatus during the translocation process. Once the Kex2 endopeptidase, encoded by the *KEX2* gene, has cleaved off the 66 residues of the pro-region, the dipeptidyl aminopeptidase A (DPAPase A, encoded by the *STE13* gene) removes both EA-dipeptides upstream of the mature terminus of the proteins. Afterwards, the mature proteins are released, the IgG assembled and finally secreted from the cell into the supernatant. **(c)** Sequence details of the junction between the αMFpp and the gene coding for the light chain. Arrows indicate the cleavage sites of Kex2 and DPAPase A. **(d)** Schematic drawing (not to scale) of the L-E and L construct for *Pichia* expression. The suffix "E" corresponds to the tetrapeptide Glu-Ala-Glu-Ala (EAEA) derived from the pro-region of the αMFpp. In the L construct the coding sequence for the EAEA-tetrapeptide was removed at the DNA level.

Surprisingly, upon analysis of purified IgGs by non-reducing SDS-PAGE (Fig. 3.2.3a), no distinct band corresponding to the full length IgG could be detected for the H / L-E and H / L constructs, both lacking the EAEA coding sequence for the heavy chain. Since this is due to a non-homogeneous composition of the heavy chain sample as will be described below, these constructs will be referred to as H* from now on. In contrast, deletion of the tetrapeptide from the light chain alone (construct H-E / L) still resulted in a clear band for the full-length IgG at the expected size. Detailed analysis of all constructs by reducing western blots (Fig. 3.2.3b) with antibodies specific for the heavy chain, for the α-factor pro sequence (αMFp) or for the lambda light chain indicated a possible explanation. In the H-E containing constructs (left four lanes in Fig. 3.2.3b, middle panel), no αMFp is detected (the light background signal is derived from the secondary antibody used which apparently cross-reacts with the human IgG), whereas in the H* containing lanes (right four lanes in Fig. 3.2.3b, middle panel) a strong signal is found. This appeared as a smear, when left untreated, but distinct bands become visible upon treatment with PNGase F. Thus, part of the heavy chain still possesses the complete α-factor pro signal sequence at an apparent deglycosylated size of ~57 kDa and thus almost 7,000 Da too high compared to the correctly processed heavy chain. The presence of the complete pro-region also explains the undetectable band for these constructs in Fig. 3a: due to the heterogeneous glycosylation of the pro-region, the IgG molecules appeared rather as a broad smear than as a single band.

The presence of the pro-region could be further confirmed by sequencing the N-termini of the respective heavy chain proteins. Edman degradation and MS analyses (Fig. 3.2.3c) clearly indicated that deleting the EAEA tetrapeptide from the heavy chain resulted in a heterogeneous sample with three major species. Next to the previously mentioned portion still having the complete pro-region attached to it (named "pro-region", blue arrow in Fig. 3.2.3d) and a minor correctly processed protein fraction named "mature" (green arrow in Fig. 3.2.3d), also a third "wrong" species was found. This latter fraction was derived from an alternative cleavage site at position 57 of the pro-region (see red arrow in Fig. 3.2.3d). These molecules possess additional nine residues of the C-terminus of the pro-region and could also be detected by SDS-PAGE, once the gels were run long enough to provide sufficient separation (lane 2 in Supplementary Data Fig. S3.2.2c). Comparing the relative signal intensities of the Edman sequencing results, the fraction containing a correctly matured amino-terminus was determined to be only approximately 20%.

We also investigated whether the processing of the heavy chain can be improved by charge engineering, i.e., making it more similar to the beginning of the light chain (Fig. 3.2.3c) (see Supplementary Data for details). However, since the charge engineering did not have the desired effect of creating homogeneous processing of the heavy chain (see Supplementary Data Fig. S3.2.2), the complete pro-region was deleted from the heavy chain construct (illustrated in Fig. 3.2.4a). The resulting construct, named "H-pro-del", has the signal sequence (pre-region) fused directly to the N-terminus of the mature heavy chain. Examination of the new H-pro-del construct in its native and deglycosylated form by Western blot (Fig. 3.2.4b) revealed that its expression resulted in homogeneous chains of the correct size, which was also subsequently confirmed by MS analysis and Edman sequencing (data not shown).



(c)

construct	measured weight [§]	theoretical weight [#]	theoretical - H_2O [$]	Δ mass [¶]	N-terminal sequence [%]	name
heavy chain (H)						
H-E	50,376	50,378	-	2	… K R ¦ E A E A Q V Q L V Q …	
H*	49,956	49,977	49,959	3	… K R ¦ Q V Q L V Q S G P G …	"mature"
	51,010	51,005	-	- 5	… A K ¦ E E G V S L E K R Q …	"wrong"
	~ 57,000 [&]	56,916	-	nd	… L A ¦ A P V N T T T E D E …	"pro-region"
light chain (L)						
L-E	22,922	22,925	-	3	… K R ¦ E A E A D I E L T Q …	
L	22,521	22,526	-	5	… K R ¦ D I E L T Q P P S V …	

[§] - mass (Da) measured by ESI-MS under reducing conditions after deglycosylation by PNGase F
[#] - mass (Da) calculated from the sequence of the unmodified, reduced mature chain
[$] - mass (Da) calculated considering pyroglutamate formation (loss of a water molecule)
[¶] - mass (Da) difference between the best matching theoretical value and the PNGase F treated measured value
[%] - the arrow indicates position of cleavage leading to the observed N-terminal sequence
[&] - mass (Da) estimated from the running behavior on SDS-PAGE under reducing conditions

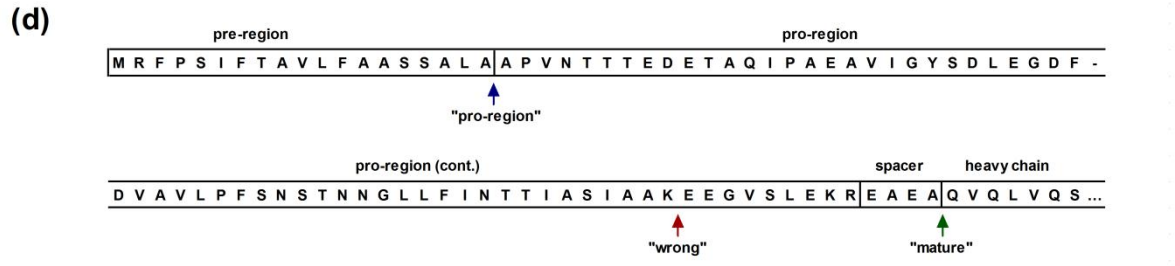


Fig. 3.2.3. Analysis of *P. pastoris* produced IgG variants. (a) SDS-PAGE analysis of equal amounts of the Protein A-purified *Pichia* constructs under non-reducing conditions stained with Coomassie Blue. The suffix "E" on the heavy (H) or light (L) chain construct denotes the tetrapeptide Glu-Ala-Glu-Ala (EAEA) derived from the pro-region of the αMFpp. As the constructs without the EAEA turned out to be heterogeneous at their N-terminus (see main text), they are marked by an asterisk (*). **(b)** Western blot analysis of IgGs after incubation in the absence or presence of PNGase F. Shown are the blots detected with antibodies specific for the heavy chain (α-H), for the α-factor pro-region (α-αMFp) and for the lambda light chain (α-L). **(c)** Overview of the measured masses of heavy and light chains. All masses are stated in Da. The penultimate column shows the corresponding N-terminal protein sequence determined by Edman sequencing. A vertical dashed line indicates the position of enzymatic cleavage within the maturation process. The color code highlights the charge of the individual amino acids (blue, positive charge; red, negative charge). The names mentioned in the last column are in agreement with those of Supplementary Data Fig. S3.2.2. **(d)** Schematic representation of the amino acid sequence of the α-factor pre-pro-region and the adjacent heavy chain protein. Arrows indicate the sites of Kex 2 cleavage, resulting in the "pro-region" variants (blue arrow), "wrong" variants (red arrow) and correct "mature" protein chains (green arrow).

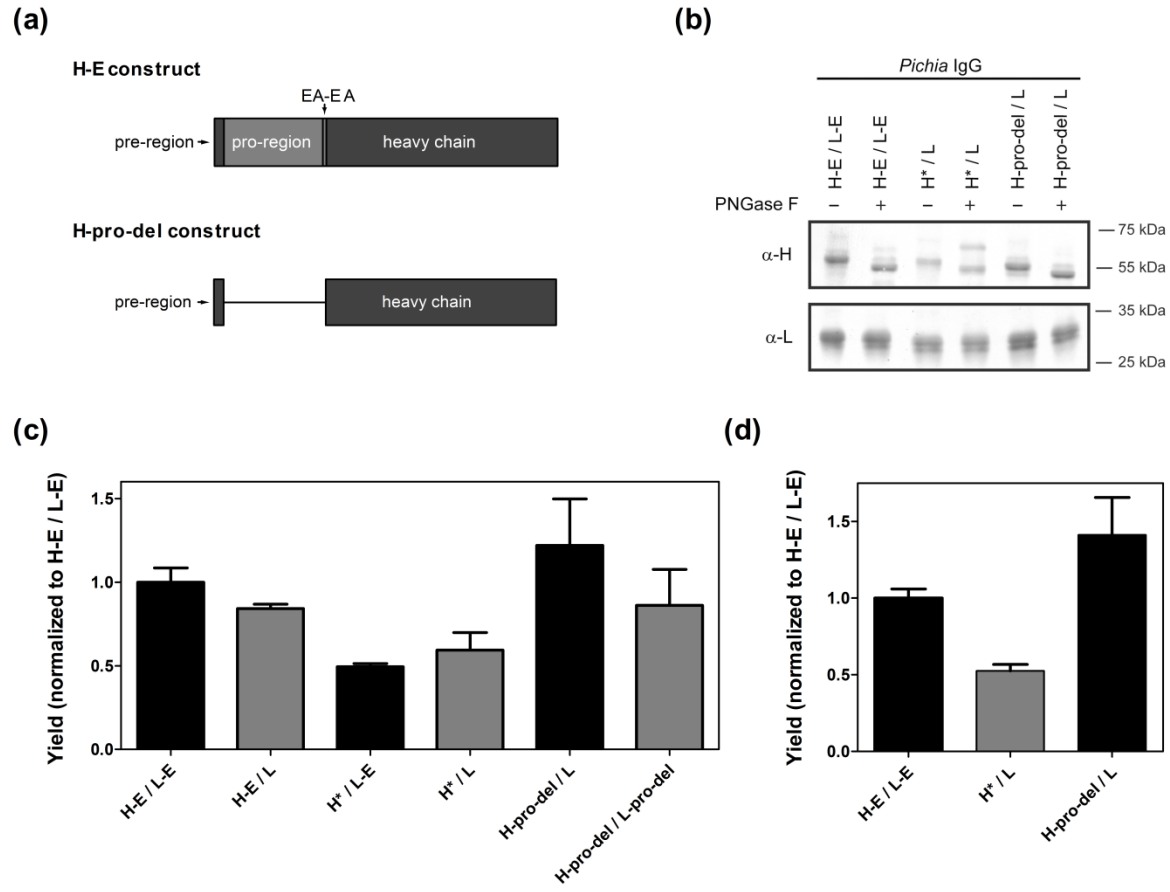


Fig. 3.2.4. Design of correctly processed IgG heavy chain in *Pichia*. (a) Schematic representation (not drawn to scale) of the H-E and H-pro-del construct for *Pichia* expression. The latter construct was designed by deleting the complete pro-region at the DNA level. **(b)** Western blot analysis of IgG constructs after incubation in the absence or presence of PNGase F. Shown are the blots detected with antibodies specific to the heavy chain (α-H) and to the lambda light chain (α-L). **(c)** Secreted IgG levels of various *Pichia* constructs, detected by sandwich ELISA in the supernatant of stable *Pichia* clones. H* indicates the incorrectly processed, heterogeneous heavy chain; H-pro-del indicates the deletion of the complete pro-region from the heavy chain construct; L-pro-del implies that this region is also deleted from the light chain. **(d)** Secretion level of corresponding Fab fragments expressed in stable *Pichia* clones analyzed by ELISA.

We also wanted to determine how the amounts of the correctly assembled IgG constructs found in the supernatant compared to all other previously described constructs (Fig. 3.2.4c). These levels could be analyzed by sandwich ELISA, capturing the secreted IgGs by antibodies specific for the heavy chain and detecting the recombinant IgG with antibodies specific for the light chain. Using this assay, it could be guaranteed that only correctly assembled IgGs were detected and, e.g., no light chain dimers affected the results. Strikingly, the heterogeneous H* constructs were found to be secreted at approximately only half the yield of the original H-E / L-E construct (49.3% for the H* / L-E and 59.3% for the H* / L construct, respectively). The new H-pro-del variant, however, seemed to overcome this limited expression and led to a total IgG amount even improved by ~25%, compared to the original H-E / L-E construct. Interestingly, the additional omission of the pro-region from the light chain (”H-pro-del / L-pro-del” construct) led to a further decrease in the IgG yield found in the supernatant of the yeast cultures.

To confirm these findings, the effects of the different pro-regions attached to the heavy chain were analyzed in the context of the corresponding Fab fragment as well (Fig. 3.2.4d). In these Fab fragments, the heavy chain V_H-C_H1 domains are connected to the complete lambda light chain by a C-terminal disulfide bond. In agreement with the data derived from the full-length IgGs, the expression of the H* / L Fab construct resulted in molecules having the complete pro-region attached at the N-terminus of the heavy chain (data not shown) and expression levels of only half the level of the original H-E / L-E variant (52.6%). The H-E / L-E level could again be increased by 41.3% by using the H-pro-del version.

Expression and analysis of mammalian IgG constructs

To examine the influence of the EAEA tetrapeptides on the aggregation susceptibilities of various IgG constructs, the corresponding molecules needed to be generated independent of the processing preferences of the expression system. Therefore, new constructs for HEK expression were designed. Since the process of protein maturation in the mammalian system is comparatively simple (Fig. 3.2.5a) different extensions could easily be attached to the amino-termini of both protein chains.

Figure 3.2.5b compares the variations made apart from the EAEA tetrapeptide. To examine whether just any random tetrapeptide might influence the biophysical characteristics of aggregation or whether the *Pichia*-produced EAEA is of special relevance, the completely unrelated and uncharged control peptide Ala-Gly-Ile-Gln (AGIQ; represented by the suffix "A") was also attached to both the heavy and light chain in different combinations. Furthermore, the light chain variant L-EA, carrying both tetrapeptides combined in the orientation depicted, was included in the study to analyze the potential importance of steric proximity of the charged residues to the original N-terminus. Various combinations of these heavy and light chain variants were created and stable HEK cell lines were established for all of them. All constructs could be expressed and resulted in proteins of expected size and banding pattern (reducing SDS-PAGE analysis in Fig. 3.2.5c; MS data not shown). To investigate the influence of the different amino-terminal extensions on the level of secreted IgGs, the amount of IgG found in the supernatant of the stable cell lines was determined by the same sandwich ELISA used for the yeast-produced IgGs. The extended mature N-terminus of the analyzed IgG chains only minimally influenced the levels of secreted IgGs (Fig. 3.2.5d). After 24 hours of expression, the level of H-E / L-E found in the media reached almost 90% of the amount derived from the original H / L construct. This is in clear contrast to the yeast αMFpp system, as detailed above.

The main biochemical and biophysical characteristics of the N-terminal extension mutants from mammalian and yeast origin were analyzed to determine potential impacts of the various tetrapeptides on the IgG structures and stabilities. The addition of the EAEA tetrapeptides was found not to influence the binding of the resulting IgGs to their antigen myoglobin (see Supplementary Data Fig. S3.2.3). The different K_D values derived from microscale thermophoresis measurements are all within the experimental error of each other.

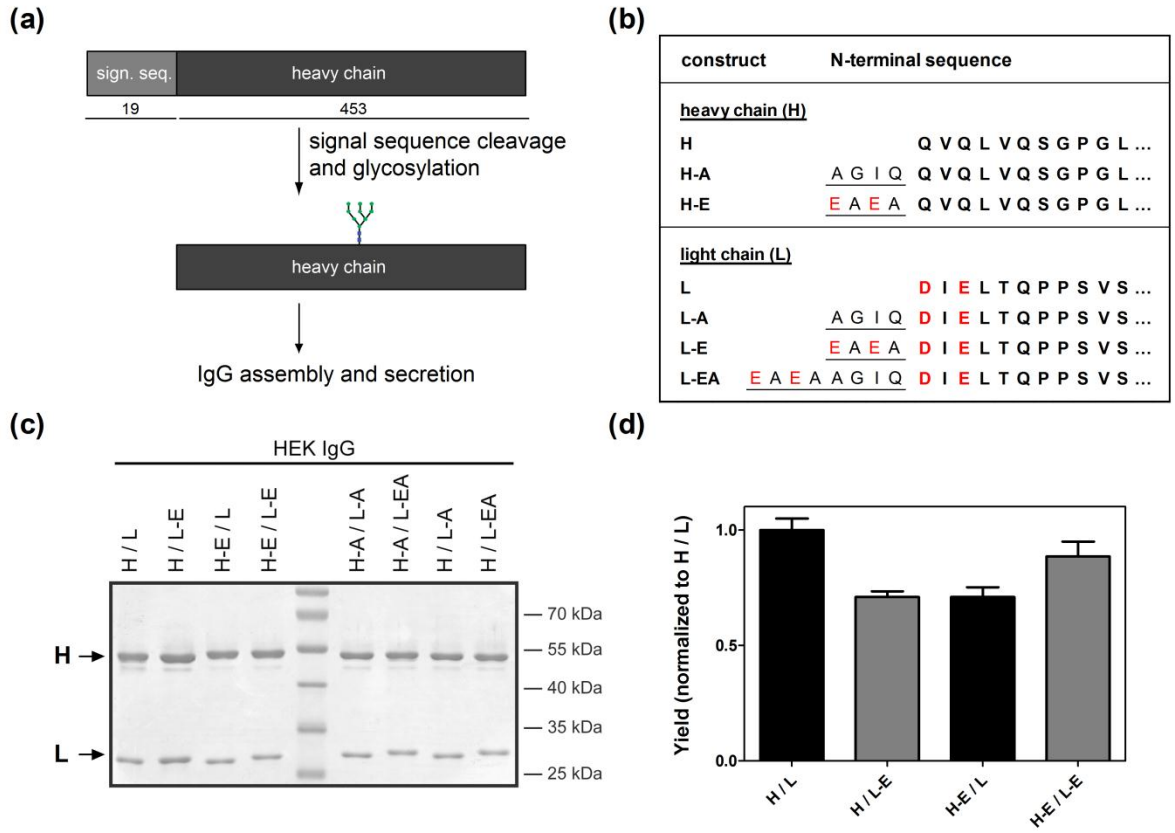


Fig. 3.2.5. Overview of different IgG constructs expressed in mammalian cells. (a) Schematic overview (not drawn to scale) of N-terminal processing of proteins in mammalian cells. Sign. seq. indicates the signal sequence. **(b)** Comparison of the mature N-terminal amino acid sequences of different mammalian heavy (H) and light (L) chain variants analyzed in this study. The suffix "A" corresponds to the control tetrapeptide Ala-Gly-Ile-Gln (AGIQ), the tetrapeptide Glu-Ala-Glu-Ala (EAEA) derived from the pro-region of the αMFpp is abbreviated as "E". The combination of both letters indicates the presence of all eight additional residues in the order presented. Red colored amino acids are negatively charged. **(c)** SDS-PAGE analysis of HEK-produced IgG variants under reducing conditions, stained by Coomassie Blue. H and L indicate the respective chains. **(d)** Secretion level of IgG variants expressed by stably transformed HEK293 cells detected by ELISA.

Figure 3.2.6 compares structural characteristics of some selected HEK-produced IgGs under native conditions and their behavior under thermal and chemically-induced unfolding. CD spectra recorded at room temperature for different mammalian variants are practically superimposable, indicating that the N-terminal extensions had no effect on the overall structure of the IgG, as expected (Fig. 3.2.6a). On the other hand, slight differences for the various E-variants were revealed in thermal denaturation experiments measured by intrinsic tryptophan fluorescence (Fig. 3.2.6b). The unfolding curves were derived from the intensity ratio of the emission spectrum at 330 nm (F_{330}) and 350 nm (F_{350}) upon excitation at 295 nm, plotted as a function of temperature.[57] One of the biggest advantages of this technique is the fact that this ratio is fairly robust against the loss of intensity due to light scattering from soluble aggregates. As can be seen in Fig. 3.2.6b, the addition of the EAEA tetrapeptide to the light chain seemed to slightly decrease the overall stability of the IgGs.

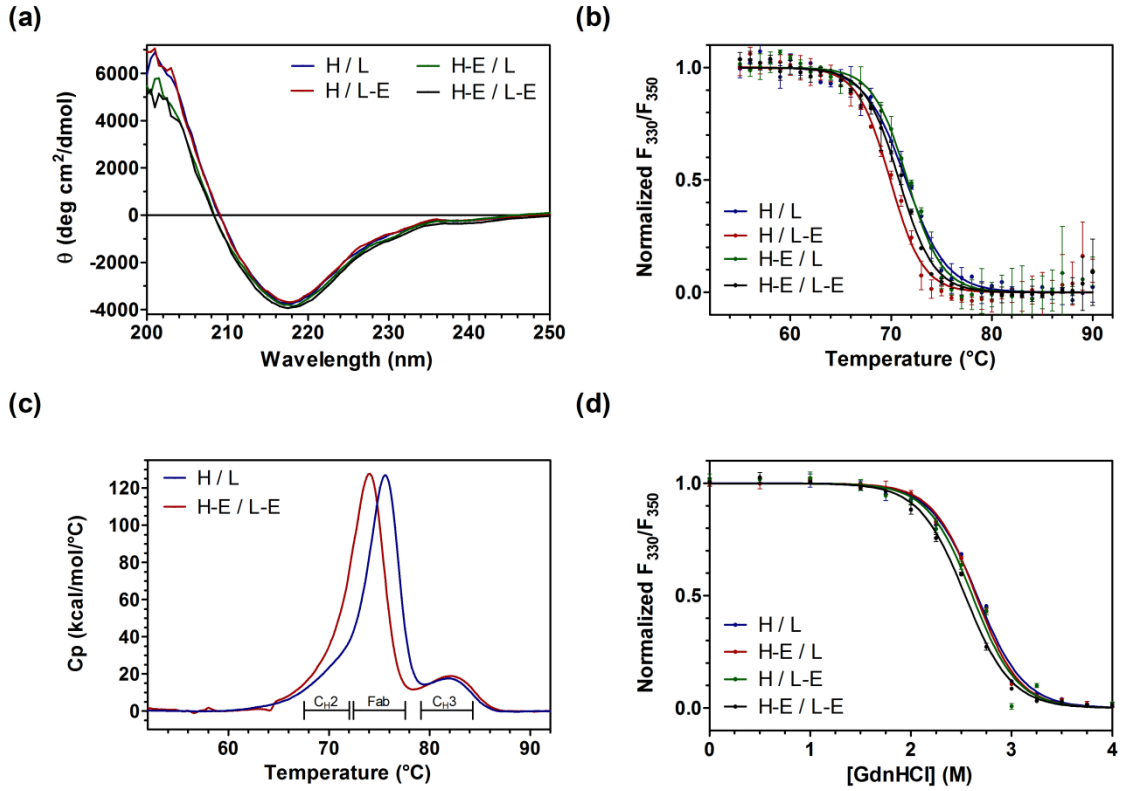


Fig. 3.2.6. Biophysical characterization of IgG constructs produced in mammalian cells. (a) CD spectra of different IgGs produced in mammalian cells. The values are reported as mean residue ellipticity (θ). **(b)** Thermal denaturation curves measured by intrinsic tryptophan fluorescence. The curves were obtained from the intensity ratio of the emission spectrum at 330 nm (F_{330}) and 350 nm (F_{350}) upon excitation at 295 nm, plotted as a function of temperature. A comparison between the different mammalian IgG variants is shown. **(c)** Differential scanning calorimetry (DSC) of the H / L and H-E / L-E construct. Indicated by the lines under the curves are the regions in which the individual domains (C_H2 domain, Fab fragment and C_H3 domain) undergo their transitions. **(d)** Guanidine hydrochloride (GdnHCl)-induced denaturation of IgGs expressed in mammalian cells. The denaturation was followed by plotting the F_{330}/F_{350} ratio as a function of increasing GdnHCl concentration.

To further investigate this finding, the H / L and the H-E / L-E constructs were examined by differential scanning calorimetry (DSC) (Fig. 3.2.6c). This method is especially suitable for distinguishing different transitions in multidomain proteins (such as IgGs) that are often “silent” or overlapping in spectroscopic methods like intrinsic tryptophan fluorescence. The previous results of a slightly decreased IgG stability could be validated, indicating that the H-E / L-E Fab fragment has its transition maximum at 74.5°C, compared to 75.8°C for H / L. These results are comparable to those obtained by guanidine hydrochloride (GdnHCl)-induced denaturation (Fig. 3.2.6d). In these experiments, the denaturation was followed by plotting the F_{330}/F_{350} ratio as a function of increasing GdnHCl concentration. These curves again suggested a slightly reduced stability for those constructs containing the EAEA tetrapeptide located at the N-terminus of the light or heavy chain. This overall tendency of a slightly reduced stability of the H-E / L-E constructs could further be confirmed by the analysis of thermal and chemically-induced unfolding of *Pichia*-produced IgGs (Supplementary Data Fig. S3.2.4).

Determination of the aggregation susceptibilities of various IgG constructs

To test whether the EAEA tetrapeptides are indeed the reason for the very different aggregation susceptibilities of HEK- and *Pichia*-produced IgGs (Fig. 3.2.1), the mammalian variants (equipped with these N-terminal extensions) were analyzed by thermal unfolding recorded by CD at 208 nm. For clarity, Figure 3.2.7 only compares the signals derived from the most important variants containing the EAEA tetrapeptides. Further data from all tested variants including the AGIQ controls can be found in Supplementary Data Fig. S3.2.5. The thermal scans provided in Fig. 3.2.7a clearly indicate a dramatic increase in the temperature of aggregation onset for those variants carrying the EAEA appendix at the N-terminus of their light chain. Interestingly, the H-E / L construct, having the identical appendix but at its heavy chain, only showed a minor increase of its aggregation onset temperature (+0.5°C). The addition of the control tetrapeptide (Supplementary Data Fig. S3.2.5) even slightly increased the aggregation susceptibility of mammalian IgGs (-0.5°C). This indicates that EAEA specifically - and not any random peptide - has this beneficial effect detailed above. Panel (a) in Supplementary Data Fig. S3.2.5 demonstrates that all constructs represented by reddish symbols have a higher temperature of aggregation onset and carry the EAEA tetrapeptide at their light chain.

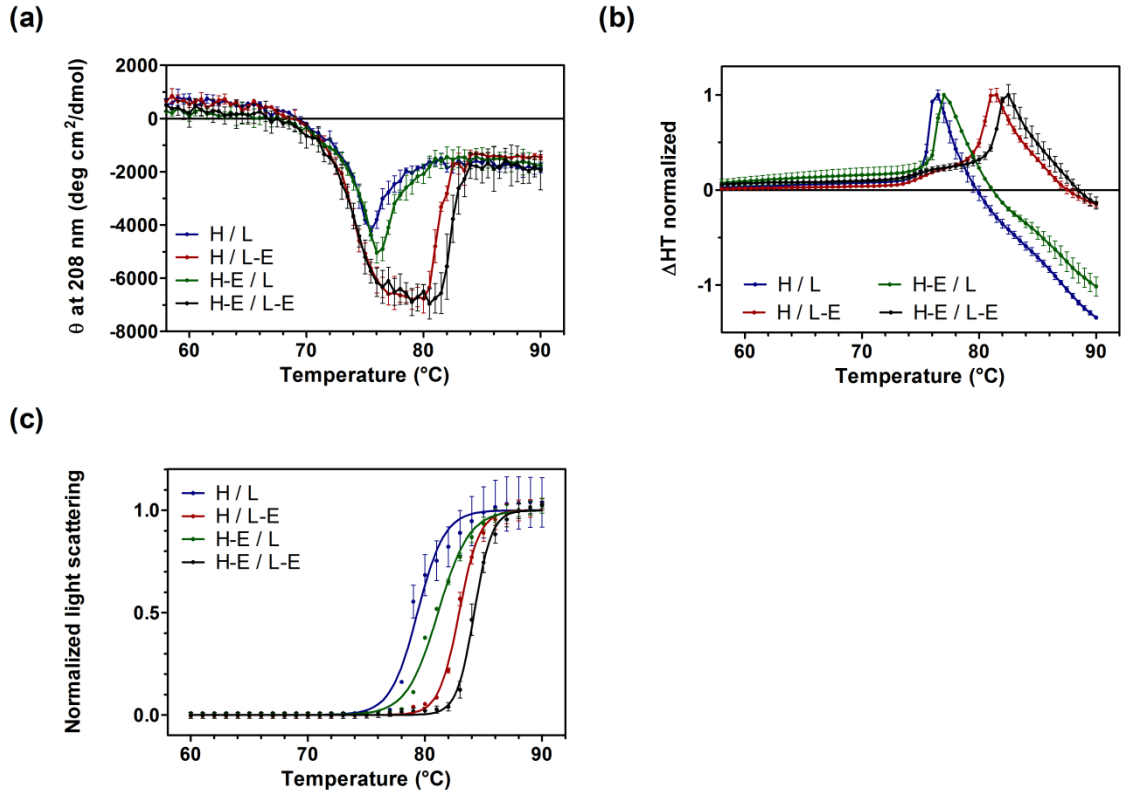


Fig. 3.2.7. Aggregation behavior of different IgGs obtained from mammalian expression. Shown are the signals derived from the most important variants containing the EAEA-tetrapeptides. Further data from additional variants can be found in Supplementary Data Fig. S3.2.5. **(a)** Thermal denaturation curves. The denaturation was followed by CD measurements, plotting the signals at 208 nm as a function of temperature. A comparison of different mammalian IgG variants is shown and the values are reported as MRE. **(b)** The high tension (HT) voltage signals corresponding to the curves shown in (a) recorded by the photomultiplier connected to the CD spectrometer are a measure of sample transparency and thus aggregation and subsequent clearance by aggregate precipitation. The values were normalized by setting the initial and the highest values of the curves as 0 and 1, respectively. **(c)** Aggregation of IgG constructs measured by light scattering at 500 nm, with excitation and emission at this wavelength.

The differences in aggregation susceptibilities could also be verified by the analysis of the signals recorded by the photomultiplier connected to the CD spectrometer denoted as high tension voltage (HT). The HT signals shown in Fig. 3.2.7b belong to the curves shown in Fig. 3.2.7a (further curves can be found in Supplementary Data Fig. S3.2.5a and b). The changes in the HT signals confirmed the formation of aggregates which induced scattering of the UV light. As this led to the reduction of the number of photons reaching the detector, the HT voltage was increased. The drop of signal intensity following the maximum was finally caused by precipitation of the aggregates. The temperatures for the maximal HT signal were determined to be 76.5°C for the H / L, 77°C for the H-E / L, 81.5°C for the H / L-E and 82.5°C for the H-E / L-E construct, respectively. These results were comparable to those derived from the aggregation assay observed in Fig. 3.2.7c. By this method, it could once more be shown that the addition of EAEA to the light chain - but not to the heavy chain - resulted in an apparent increase in the temperature of aggregation onset for the H / L-E construct by ~5°C, compared to the H / L variant.

The results obtained from the mammalian cell-produced IgGs could be confirmed by the analysis of their counterparts produced in the yeast *Pichia pastoris*. Supplementary Data Fig. S3.2.6 shows the analysis of the aggregation behavior of various *Pichia*-produced IgGs, analyzed by CD spectroscopy and the aggregation assay. In agreement with the data derived from HEK IgGs, the presence of the EAEA tetrapeptide in the yeast sequences resulted in IgG variants with decreased aggregation tendencies - emphasizing that the effect of this peptide on aggregation susceptibility required its location specifically at the N-terminus of the light chain.

To investigate the effect of the four extraneous amino acids at the light chain under more physiological conditions, the formation of aggregates after extended incubation of freshly prepared IgGs was measured by static multi-angle light scattering (MALS), coupled to size exclusion chromatography (SEC). Panels (a) and (b) in Figure 3.2.8 display the normalized light scattering measured for the HEK-expressed H / L and H / L-E variants, respectively, after a five day incubation at either 4°C, 37°C or 50°C. Incubation at elevated temperatures dramatically increased the light scattering of the control IgG (H / L) for an aggregate peak eluting at ~8 ml (panel (a)). This larger signal was due to both an increased amount as well as the larger size of these aggregates. Consequently, the monomeric peak at ~13 ml decreased in its intensity. This was, however, also partly due to the disassembly of both light chains from the antibody, detectable as a little shoulder towards higher elution volumes and in SDS-PAGE (data not shown). Quite different results were obtained once the EAEA tetrapeptide was attached to the light chain. As depicted in panel (b), the amount (i.e. intensity) of the aggregation peak did hardly change even under more stressful conditions (50°C) and only a minimal increase in the size of aggregates (i.e. a shift towards smaller elution volumes) was detected. As for H / L, the monomeric IgG peak, however, still was slightly diminished due to IgG disassembly.

To get further insights into the effect of the EAEA peptide on long-term stability, IgG samples were also incubated for six months at 4°C and afterwards run on an analytical SEC as illustrated in Fig. 3.2.8c. Although only a small fraction of the sample turned into aggregates (and eluted at ~0.9 ml), a clear difference in intensity could be seen for the various constructs. While the H / L variant exhibited

an aggregated fraction of almost 3%, this peak was noticeably decreased for the H-E / L-E construct, adding up to only 0.5% of the total protein.

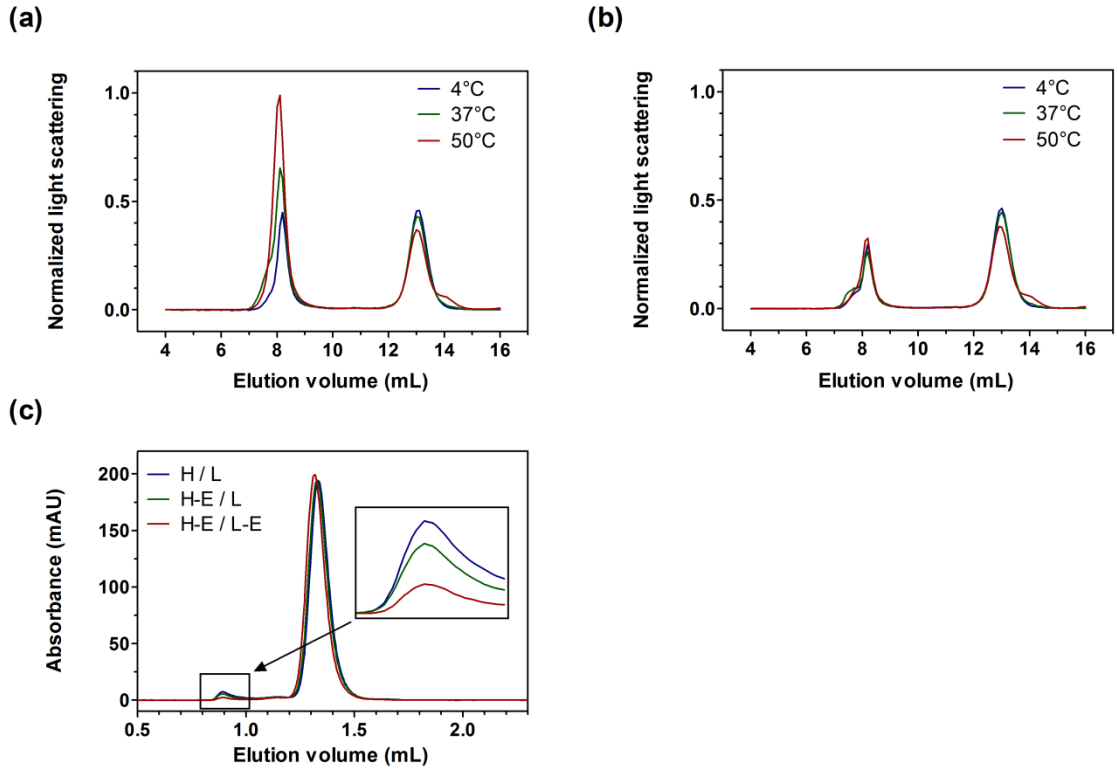


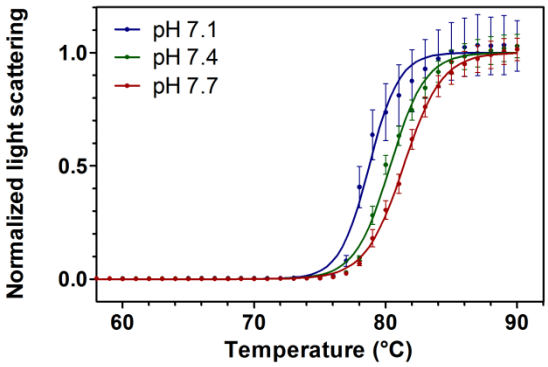
Fig. 3.2.8. Aggregation of different IgGs upon incubation at 4°C, 37°C or 50°C analyzed by MALS and SEC. (a) Light scattering of the IgG H / L construct, analyzed by multi-angle light scattering (MALS), coupled to size exclusion chromatography (SEC). Samples were previously incubated for five days at 4°C, 37°C or 50°C, respectively. **(b)** Light scattering of the IgG H / L-E construct. Plotted are the signals obtained from samples incubated for five days at 4°C, 37°C or 50°C, respectively. **(c)** SEC of various HEK-produced IgG constructs after storage at 4°C for six months. Note that (a) and (b) were done with a 24 ml Superdex 200 10/300 GL column, while in (c) a 2.4 ml Superdex 200 PC 3.2/30 column was used.

Influence of pH values on the temperature of aggregation onset

Since the difference between the EAEA peptide and the ineffective control peptide AGIQ are two negative charges, the influence of pH values on the aggregation susceptibilities was examined more closely. As depicted in Fig. 3.2.9a, even minor changes of the pH value dramatically influenced the temperatures of aggregation onset. While the H / L-A construct aggregated at 78.8 ± 0.1°C at the same pH as in all previous analyses (pH 7.1), changing it to pH 7.4 increased this temperature to 80.3 ± 0.1°C, and even further to 81.4 ± 0.1°C in PBS adjusted to pH 7.7. Using a web-based protein calculator (www.scripps.edu/ ~cdputnam/protcalc.html) the net charge of the different IgG constructs at various pH values was computed. In the second column of Fig. 3.2.9b, the net charges of full-length IgG consisting of both heavy and light chains are shown for various pH values. Adding the EAEA extension to the light chain (third column) had the same effect on the protein's overall charge as increasing the pH by 0.35 units, namely adding four negative charges to the molecule (two for each light chain). This addition of negative charge had such a dramatic effect on aggregation susceptibility since the charge of the H / L construct is minimally positive at the pH values generally used throughout

this study (pH 7.1). Therefore, lowering the overall charge by four units has a rather profound effect on the net charge of the whole molecule.

(a)



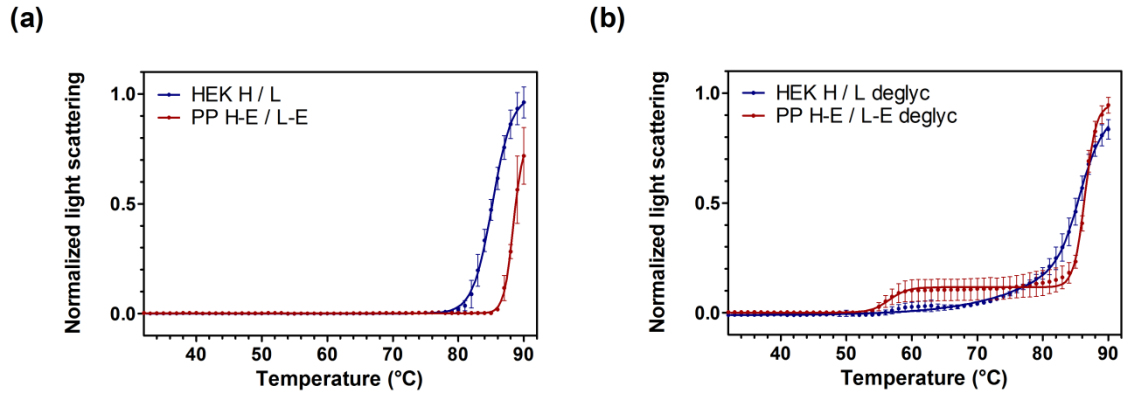
(b)

pH	charge of constructs		
	H / L	H / L-E	λ-L
7.0	0.4	- 3.6	- 6.7
7.1	- 0.9	- 4.9	- 6.9
7.2	- 2.0	- 6.0	- 7.0
7.3	- 3.2	- 7.2	- 7.2
7.4	- 4.3	- 8.3	- 7.4
7.5	- 5.5	- 9.5	- 7.6
7.6	- 6.7	- 10.7	- 7.8
7.7	- 8.1	- 12.1	- 8.0

Fig. 3.2.9. Influence of the pH on the aggregation of IgG. (a) Aggregation of IgG H / L-A construct measured by light scattering at 500 nm. Samples were previously equilibrated in PBS buffer at the indicated pH. **(b)** Overview of the overall charge of the different IgG constructs at various pH values. Analyses were performed using the web-based protein calculator (www.scripps.edu/~cdputnam/protcalc.html). In the first three columns, the results for the full-length IgG consisting of two heavy and light chains (including indicated N-terminal appendices) are shown. Two additional negative charges have been considered due to the formation of N-terminal pyroglutamate residues at the heavy chain, thus lacking the positively charged N-terminal amino group. The last column specifies the charges derived from analyzing only one light chain.

Transfer of findings onto a second IgG construct

Finally, we analyzed to which extent our findings could be transferred to other IgG constructs. Since the effect seemed to be related to the overall charge of the molecule, we chose a completely unrelated IgG as an alternative model system. Having very different CDR regions in its kappa light chain and thereby a comparatively high isoelectric point (pI) of 6.9 (the previously used lambda light chain has a pI of 4.9 due to its acidic CDRs), this construct had a completely different range of overall IgG charges. When tested in the aggregation assay (Fig. 3.2.10a), the yeast-produced version with its different glycan moiety indeed aggregated at higher temperatures compared to its mammalian counterparts (88.4 ± 0.1°C vs. 82.2 ± 0.1°C), indicating that this glycan effect seems to be general. However, once the glycans were removed by PNGase F treatment, only minor differences in the temperature of aggregation onset were observed (Fig. 3.2.10b; 86.4 ± 0.1°C vs. 85.5 ± 0.3°C). The reason for this might be that the overall charge of this new IgG construct is more positive than for the first analyzed IgG, such that the addition of the EAEA tetrapeptide has a less profound influence on the overall charge in this case (Fig. 3.2.10c).



(c)

pH	charge of constructs		
	H / L	H / L-E	κ-L
7.0	20.8	16.8	0.3
7.1	19.5	15.5	0.1
7.2	18.3	14.3	0.0
7.3	17.1	13.1	- 0.2
7.4	15.9	11.9	- 0.4
7.5	14.6	10.6	- 0.6
7.6	13.4	9.4	- 0.8
7.7	12.0	8.0	- 1.0

Fig. 3.2.10. Aggregation behavior of a second, kappa-light chain containing IgG. Comparison of a different IgG expressed in either *P. pastoris* or mammalian cells. The mammalian construct is of the H / L design whereas the IgGs from the *Pichia* system are still of the H-E / L-E format. Shown are the light scattering data of **(a)** the original and **(b)** the PNGase F-deglycosylated IgGs. The additionally detectable light scattering signal starting at ~56°C is due to the unfolding and aggregation of the PNGase F enzyme. **(c)** Overview of the overall charge of the second IgG constructs at various pH values. Analyses were performed using the web-based protein calculator (www.scripps.edu/~cdputnam/protcalc.html) (as in Fig. 3.2.9b). Two additional negative charges have been considered due to the formation of N-terminal pyroglutamate residues at the heavy chain, lacking the positively charged N-terminal amino group.

Discussion

When immunoglobulin G molecules of identical amino acid sequence were produced in different eukaryotic systems, dramatic differences in their aggregation susceptibilities were encountered. This observation turned out to be due to two major factors: (i) differences in the glycan structures and (ii) residues remaining from the *Pichia* secretion system used, the α-factor pre-pro sequence (αMFpp).

Glycan structures

Antibodies of the IgG format are N-glycosylated at the Asn297 position in their C_H2 domains. These domains, having the lowest stability within the IgG molecule,[58] are the only ones not having any direct protein-protein interaction and solely contact each other by their sugar moieties. The attached glycans have proven to fulfill multiple roles in the maturation and function of IgGs. In general, glycosylation is a means of "quality control" in the eukaryotic ER.[59; 60] Indeed, the glycosylation of the antibody's C_H2 domain has recently been shown to be a critical step in antibody folding.[61; 62]

In the mammalian IgG molecule, the sugars have further taken on the functional role of enabling the antibody to bind to the activating Fcγ receptors.[18] This has been interpreted as the sugars mediating both the distance and the relative mobility of the domains towards each other, which is necessary since Fcγ receptors bind the C_H2 domains near the hinge region.[63-66] Even one single fucose residue can decrease the affinity of IgG for FcγRIII as it sterically interferes with the receptor's own oligosaccharide,[67] while fucosylated IgGs still bind to the inhibitory FcγRII equally well, having important consequences for glycoengineering.

Nonetheless, an increasing number of therapeutic programs are not aiming for FcγR activation, but indeed actively avoiding it.[68] Thus, it is worthwhile to consider alternative expression systems for IgG production, including the yeast *Pichia pastoris*. For this production system, however, the role of glycosylation for the stability of the native IgG under storage and *in vivo* conditions has to be considered. When comparing IgGs with complex glycosylation from HEK293 cells to those containing *Pichia*-produced mannose-rich glycosylation, we found two decisive differences. First, the melting temperature of the corresponding C_H2 domain and thereby of the whole IgG is lowered when produced in *Pichia*, as can be seen by three independent techniques: circular dichroism (Fig. 3.2.1a), intrinsic tryptophan fluorescence (Supplementary Data Fig. S3.2.1f) and differential scanning fluorimetry (DSF, Supplementary Data Fig. S3.2.1d). In the latter technique, which like DSC offers increased resolution of the thermal unfolding of different domains, the shift in the stability of the C_H2 domain to lower temperatures due to the *Pichia* glycosylation is clearly visible. The second difference, however, is more than counterbalancing this negative effect: As detailed above, IgGs with *Pichia* glycosylation are much less prone to aggregation.

In agreement with our data, *Pichia* glycans carry more sugar units and are known to be rich in mannose units,[22] which may sterically block the initial docking of IgGs to each other at the onset of aggregation. It is also known that these oligosaccharides may contain several mannosylphosphates, transferred to both the core and outer sugar chains.[21; 22; 69; 70] However, our analysis showed that our IgGs produced in *Pichia pastoris* carried no negative charge - and thus no mannosylphosphate - in their oligosaccharides. We arrived at this conclusion by isoelectric focusing (IEF) analysis of the IgGs. It revealed that the isoelectric points (pI) of both *Pichia* and HEK produced antibodies before and after deglycosylation by PNGase F were almost identical (see Supplementary Data Fig. S3.2.7). If a negative charge had been removed from the IgGs produced in *Pichia*, this would have been seen as a shift upwards to less negative pI upon deglycosylation. In addition, the running behavior of *Pichia* produced IgGs on ion exchange chromatography did not change upon deglycosylation (data not shown) – an effect that would have been expected if a certain fraction of the resulting *Pichia* glycans had possessed negative charges on their mannose units. Therefore, the beneficial effect of the *Pichia* glycans conferring resistance to aggregation does not seem to be based on any glycan contribution to the protein's overall charge.

To highlight the importance of the glycan moiety, we have included unglycosylated T299A IgG constructs in our studies, confirming the role of the sugars for stability and in protection against aggregation (Supplementary Data Fig. S3.2.1 and Fig. 3.2.1c). Unglycosylated IgGs melt at lower temperatures and are more aggregation prone than IgGs glycosylated with either *Pichia* or mammalian

oligosaccharides - a finding consistent with previous observations.[71] Fig. 3.2.11 compares the aggregate formation for IgGs with identical primary sequences and N-terminal extensions (H-E / L-E constructs) made in either expression system. Upon deglycosylation by PNGase F, the molecules should be of identical composition, whether expressed in either mammalian or *Pichia* cells, except for the removal of the C-terminal lysine from the heavy chains in mammalian cells, which was not observed to occur for the *Pichia* constructs. Indeed, the curves indicating the aggregation behavior are almost identical.

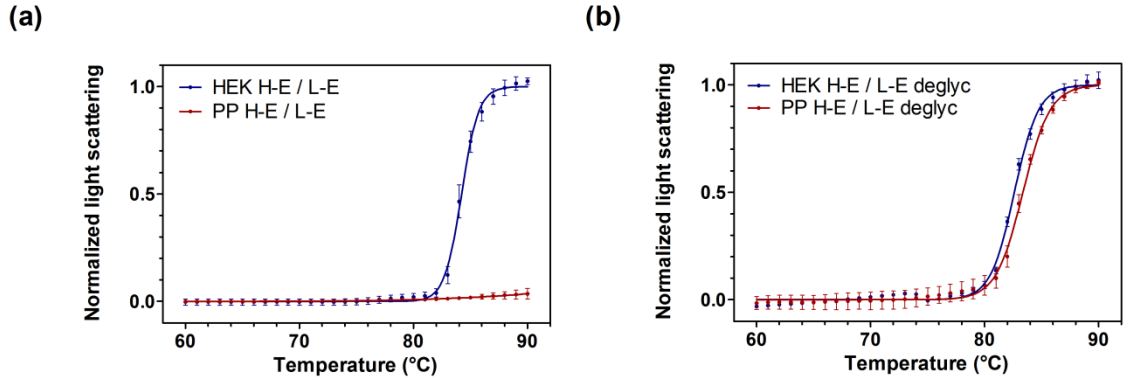


Fig. 3.2.11. Aggregation behavior of the same IgG constructs produced in different mammalian expression systems. (a) Aggregation of the IgG H-E / L-E construct produced in either *P. pastoris* (PP) or mammalian cells (HEK293). Having identical amino acid sequences, these samples only differ with respect to their post-translational modifications such as glycosylation. Aggregation was measured by light scattering at 500 nm and plotted as a function of temperature. **(b)** Aggregation of the same IgGs as in (a) in their deglycosylated state. Samples were treated by PNGase F prior to analysis by light scattering.

Taken together, the *Pichia* glycosylation has the surprising effect of protecting the IgG against aggregation better than the mammalian complex sugar moiety. For future applications, this must of course be weighed against the therapeutic necessity of FcγR activation. Recently, glycoengineered *Pichia* strains have been developed and meanwhile successfully applied to the production of monoclonal antibodies, such that options for both types of glycosylation exist.[23-26]

Processing in *Pichia pastoris*

Most secreted recombinant proteins produced in *P. pastoris* are constructed as fusions to the α-factor pre-pro sequence (αMFpp) derived from *S. cerevisiae*.[72] The α-factor peptide itself is separated by a EAEA spacer from the pro-sequence, necessitating a two-step cleavage by Kex2 and DPAPaseA.[73] Next to making the Kex2 processing site accessible, the spacer could also help to keep the resulting product more aggregation-resistant by providing charged extensions. While the natural α-factor peptide is soluble, the precursor contains several copies of the mature peptide, and thus it is conceivable that some intermediate processing products will benefit from increased solubility.

Once a recombinant protein is to be secreted with this system, the question arises whether this spacer should be included or not. It has previously been shown that its inclusion can increase the amount of correctly processed material by preventing steric hindrance of the Kex2 cleavage site.[74] Whether this is necessary for the specific POI may depend on its primary sequence around the cleavage site,[75] its

structure and thus the accessibility to Kex2, and potentially its aggregation tendency, as postulated in this study. While the inclusion of the spacer generally enhances the cleavage efficiency of this endopeptidase specific for dibasic sites, several studies have demonstrated problems with the subsequent proteolytic processing of the EA repeats.[54-56; 76] Apparently, the quantity and / or activity of the *STE13*-encoded dipeptidyl aminopeptidase A in the secretory pathway is not sufficient to process these repeats for large amounts of recombinant protein expressed under strong promoters. The resulting secretion of molecules with amino-terminal EAEA extensions was also observed for both the heavy and the light chain of the initial IgG constructs expressed in the *Pichia* system in this study.

Since the presence of these amino-terminal extensions can be undesirable for many applications and even might create a new immunogenic epitope,[77] the omission of the EAEA tetrapeptides was evaluated, aiming to achieve authentic N-termini at both the heavy and light chain. This simple strategy worked well for the expression of the light chain components, leading to proteins of the correct sequence, while causing only a slight reduction of the total amount of secreted IgGs (Fig. 3.2.4c: H-E / L-E vs. H-E / L).

However, deleting the EA dipeptides from the heavy chain constructs caused severe problems in its maturation process. Three species of resulting heavy chain proteins were observed: (i) a fraction containing the complete and unprocessed pro-region, (ii) molecules containing the last nine amino acids of the pro-region on its N-terminus (resulting from incorrect processing), and (iii) the correctly processed heavy chain. Similar results were achieved when analogous changes were carried out with Fab fragments (data not shown). The underlying decreased Kex2-cleavage efficiency in the absence of a spacer is in agreement with previous studies reporting that expression of proteins without the EAEA spacer peptide led to the secretion of incompletely or incorrectly processed proteins.[78-80] Remaining precursor was observed when onconase was fused to the αMFpp without the EAEA spacer.[81] The problem of uncleaved pro-regions seems not to be specific for the *Pichia* system, as similar findings had already been observed earlier for the expression of an insulin precursor in *S. cerevisiae* with the αMFpp system in the absence of any EAEA spacer.[82]

The presence of an additional cleavage site within the pro-region was also previously described for a different single-chain antibody in V_H-linker-V_L orientation.[55] However, in that case, eleven residues from the carboxy-terminus of the pro-region were found to be attached to the protein, resulting from a cleavage at position 57 of the pro-region between ..-I-A‖A-K... (where ‖ indicates the Kex2 cleavage site). In our case, this cleavage was shown to have occurred between ..-A-K ‖ E-E-.. and thereby two residues further downstream in the pro-region, resulting in an extension of only nine residues. This site actually seems to be more reasonable for Kex2 cleavage, taking into account that the "original" Kex2 cleavage site (..-K-R ‖ E-A-..) has a similar charge distribution, possessing a basic amino acid in the P1 position just upstream of the cleavage site and an acidic residue in the adjacent P1' position. While general DPAPase A processing can probably be increased by just overexpressing the enzyme, it would be expected that this incorrect cleavage would remain as a side reaction, and thus DPAPase A overexpression might not lead to homogeneous product either.

Secretion without the pre-pro-α-factor pro sequence

In principle, the signal sequence alone is sufficient for antibody secretion to the medium in eukaryotic cells and the pro-sequence of the α-factor is not needed. The role of the pro-region for efficient secretion has been investigated before for other proteins like the lipase Lip1p.[75] This protein could be expressed and secreted either under the direction of the α-factor pre or pre-pro sequence - however, the highest level of secretion was achieved by the construct with the full pre-pro sequence. In contrast, here the deletion of the complete pro-region from the heavy chain did not only lead to a product with the correct N-terminus, but also increased the amount of correctly assembled antibody found in the cell supernatant by 25% or 40% for the IgG or Fab fragment, respectively (Fig. 3.2.4c and d). While this study revealed that deleting the pro region is the only option for the heavy chain to obtain homogeneous protein with the genuine N-terminus, correctly processed light chain can be produced either with the pro-region (in the absence of EAEA) or even in the absence of the pro-region. Yet, for the light chain, the deletion of the pro-region resulted in a reduction of total protein yield by ~30% ("H-pro-del / L" vs. "H-pro-del / L-pro-del" constructs, see Fig. 3.2.4c). Thus, the highest yield of genuine IgG without N-terminal extensions can be obtained after deletion of the complete pro-region for the heavy chain, and the presence of the pro-region but absence of the EAEA tetrapeptide for the light chain (Fig. 3.2.4c).

The tetrapeptide extensions influence protein aggregation but not stability

To analyze the impact of various amino-terminal extensions on the biophysical characteristics of the IgGs, the ability to attach any extension to the N-terminus of any chain was crucial. Therefore, these constructs were expressed in HEK293 cells. Analysis of the sequences of the new constructs (Fig. 3.2.5b) by the signal sequence prediction software SignalP 3.0[83] (www.cbs.dtu.dk/services/SignalP) predicted that all variants should be processed correctly. Indeed, all constructs could be expressed and were confirmed to be homogeneous and in agreement with their design by mass spectrometry, SDS PAGE and sandwich ELISA. Nonetheless, compared to the H / L construct, slightly lower levels of IgG were found in the supernatant (~70% for the H / L-E and H-E / L constructs and 88% for the H-E / L-E variant, respectively) (Fig. 3.2.5d). Further investigations will have to clarify which step in biosynthesis is responsible for these differences.

Once stable cell lines were established for all mammalian constructs, their biophysical properties could be analyzed in detail. As expected, the amino-terminal extensions of either the heavy or the light chain - individually or combined in the same construct - only very slightly influenced these characteristics (Fig. 3.2.6). Only a slight decrease of ~1°C was encountered for H-E / L-E compared to the original H / L construct in the Fab transition within the IgG in differential scanning calorimetry (Fig. 3.2.6c, Supplementary Data Fig. S3.2.4). A dramatic difference, however, became apparent for the aggregation susceptibilities of the different constructs. As depicted in Fig. 3.2.7, the presence of the amino-terminal EAEA tetrapeptide at either N-terminus generally enhanced the temperature of aggregation onset. Strikingly, this increase was more pronounced for the EAEA extension on the light chain. From a comparison with the extension with the AGIQ control tetrapeptide (Supplementary Fig.

S3.2.4), these findings indicate that the two negative charges of the EAEA extension could be partially responsible for the reduced susceptibility of the resulting IgG molecules towards aggregation.

The influence of charge and location on the effect of the EAEA extension

Our data indicate that the EAEA tetrapeptide extension has more influence on preventing aggregation once it is added to the light chain rather than to the heavy chain of the antibody studied (Fig. 3.2.7c). One obvious difference between the two chains is the presence of pyroglutamate as the N-terminal residue of the heavy chain produced in mammalian cells. MS data and the results of Edman sequencing indicated that the encoded glutamine is indeed converted into pyroglutamate, as commonly seen for IgGs in general.[51] Since pyroglutamate formation from glutamine results in the loss of the positively charged N-terminal amino group, it slightly lowers the isoelectric point of the antibody.

If any tetrapeptide (either AGIQ or EAEA) is N-terminally attached to the heavy chain, the cyclization reaction forming pyroglutamate cannot take place anymore and the new construct will carry this original positive charge at its glutamine residue again plus any charges introduced by the peptide side chains. Therefore, the addition of AGIQ adds an additional positive charge to the IgG, while the addition of the EAEA tetrapeptide to the heavy chain introduces only one new net negative charge (the second one is canceled out by the positive charge of the free N-terminal amino group). Interestingly, the newly N-terminally located glutamic acid of the EAEA tetrapeptide was not found to have undergone the cyclization reaction into pyroglutamate (data not shown), even though the formation of pyroglutamate from N-terminal glutamate should in principle be possible as well.[84]

When the EAEA extension is present on the light chain, which carries an N-terminal amino group, both its negative charges are added to the protein's net charge. Taking all findings together, it seems that the EAEA tetrapeptide with its two negative charges is beneficial for the IgG’s resistance to aggregation. In general, its exact location within the molecule might per se not be of great importance, as the H / L-EA construct (where the EAEA extension is spaced by the AGIQ tetrapeptide from the N-terminus of the light chain) also exhibited an increased aggregation resistance. However, the loss of the pyroglutamate residue might reduce the effect of this extension at the heavy chain.

Our hypothesis that the positive impact of the EAEA tetrapeptide is very likely based on its effect on the overall charge of the antibody molecule was further corroborated by the observation that the exact temperature of aggregation onset depends on the pH of the sample (Fig. 3.2.9). As the charge of the H / L construct is almost neutral at the pH value generally used throughout this study (pH 7.1), this sequence extension has a rather dramatic effect, causing the whole molecule to possess a clearly negative net charge. The only marginally negative charge of the original constructs at pH 7.1 originates from the nature of their CDR regions. While the isoelectric points (pI) of the backbones are 8.2 and 5.5 for the heavy and light chain, respectively (determined by the analysis tool on the ExPASy server http://web.expasy.org/compute_pi), the basic pI of the CDRs within the heavy chain are "compensated" by a very acidic pI of the light chain’s CDRs (9.2 and 4.1, respectively). The reason for the pH affecting mostly the net charge of the heavy chain lies in its larger number of histidine groups.

Application aspects of these findings

The beneficial effect of the EAEA tetrapeptide on the aggregation susceptibilities of the analyzed antibodies was not only seen when the IgGs were analyzed by gradual heating. When freshly produced IgG variants were incubated for five days at various temperatures, clear differences in the formation of aggregates could be detected by MALS analyses (Fig. 3.2.8). Aggregate accumulation was seen for the H / L construct, while for the EAEA containing H / L-E construct neither the amount nor the size of aggregated IgGs noticeably increased. Also, when the antibodies were stored at 4°C for several months and subsequently analyzed by SEC, the H-E / L-E construct possessed the smallest aggregated fraction. The percentages derived from comparing this fraction's aggregates with that of H / L aggregates (0.5% vs. 3%) might sound insignificant and negligible - however, it should be kept in mind that the samples were stored at a relatively low protein concentration of 1 mg/ml. Since antibodies administered in the clinic are generally stored and used at much higher concentrations, this small difference might then increase dramatically.

To transfer the results of this study to other antibodies, it is important to understand to what degree the findings can be generally applied. Using another unrelated IgG, whose pI is substantially higher due to its CDR regions (light chain alone: 6.9 vs. 4.8; complete IgG: 8.9 vs. 7.2) and which is thereby more positively charged at pH 7.1, we found that there was only a small influence of the EAEA tetrapeptide (Fig. 3.2.10b). To test these effects, the IgGs were produced in HEK293 cells (which leads to correct processing to the genuine N-terminus) or in *Pichia pastoris*, where the EAEA tetrapeptide was still attached on both chains, and subsequently analyzed in the glycosylated and deglycosylated state. Compared to the previously analyzed antibody, this IgG has a substantial positive net charge at neutral pH (Fig. 3.2.10c), and even though its charge is diminished by the EAEA extension, this has only a very small effect. It thus implies that the EAEA extension will be most promising for antibodies whose net charge can be boosted significantly by this extension.

The results obtained from the second IgG, however, also indicated the transferability of the other finding of this study. The glycosylated *Pichia*-produced IgG generally showed a higher resistance to aggregation compared to its counterpart from mammalian cells. This result suggests that the improved aggregation protection by the mannose-rich oligosaccharide is seen in this antibody too, despite its high isoelectric point. It is thus likely to be a general phenomenon that might be transferable to other molecules as well.

Taken together, the present finding indicate that the convenient production in the host *Pichia pastoris* has an unexpected benefit, as it leads to antibodies being more resistant to aggregation due to both the *Pichia* glycosylation as well as the residual tetrapeptides left by the α-factor pre-pro secretion system. Importantly, through these modifications, aggregation of the IgG was also prevented under physiological conditions, such as after an extended time at 37°C. Nonetheless, the magnitude of these effects will be antibody-dependent, with the net charge seemingly being a good predictor.

Comparison with other proteins

The present findings are in agreement with previous analyses showing that the solubility of proteins can be enhanced by the introduction of charged residues resulting in an altered overall protein charge.[85; 86] In addition, our results match previous studies showing that single V_H domains that are by themselves resistant to aggregation generally possess a greater negative net charge than their aggregation-prone counterparts.[86-88] This correlation has, however, never been reported for full-lengths IgGs so far. A similar connection between pI values and aggregation susceptibility has been presumed for single-domain camelid V_{HH} antibodies. Although possessing partially lower conformational stabilities compared to their human V_H counterparts, these molecules mainly unfold reversibly and resist aggregation.[42; 89] Following a similar reasoning, a human V_H domain prone to aggregation could be converted into an aggregation-resistant variant by introducing negatively charged residues within or near the CDR1 loop.[88] However, in the same study it was shown that other negatively charged mutations in the most solvent-exposed residues outside of the CDRs were incapable of suppressing the parent V_H's aggregation propensity. Therefore, the authors concluded that the simple difference in the net charge of this domain antibody might not be responsible for the improvements of the aggregation propensities. Instead, the specific locations of the mutations seemed to be of great importance. Our results, however, indicate that an N-terminal placement of negatively charged residues might be sufficient to increase aggregation resistance. This strategy circumvents the concern that the introduction of charged residues into the CDR regions may abolish the antibody-antigen recognition directly or indirectly by altering the loop conformation.[88]

Generally, it appears that the N-terminal location is particularly suitable for the addition of charges, as it has diminished the aggregation of a molecule as large as an IgG. Compared to the previously mentioned study by Perchiacca and colleagues[88], the strategy resulting from our results should be more easily applicable to a large variety of molecules, as the charged EAEA extension can readily be attached to any molecule by simple cloning. In addition, there is no risk that these new residues might disturb the structure of the IgG as they are not inserted within the domains. Importantly, the present study could confirm that the various amino-terminal extensions did not affect the binding of the IgG variants to their antigen. As depicted in Supplementary Data Fig. S3.2.3, these interactions were analyzed for several mammalian constructs and the determined affinities were unaltered within experimental error.

Conclusions

Taken together, our results indicate that the production of full-length IgG in the yeast *Pichia pastoris* - being attractive for large scale preparations due to its convenient handling - is possible and results in molecules with the desired antigen-binding and thermal stability. However, obtaining correctly processed IgGs requires the direct fusion of a signal sequence to the heavy chain. When using the commonly utilized α-factor pre-pro secretion system, the secreted antibodies will contain the EAEA extension at both N-termini. Removing this extension from the coding sequence can easily be performed for the light chain, but will generate heterogeneous processing for the heavy chain. Most

surprisingly, antibodies produced in *Pichia pastoris* proved to be more resistant to aggregation than their counterparts with identical primary sequences produced in mammalian cells. This was shown to be caused by the mannose-rich glycosylation and the residual EAEA extensions. While the former effect appears to be general, the latter is most pronounced when this tetrapeptide addition can significantly increase the net negative charge of the antibody. Furthermore, these results show that the addition of a mere four amino acids to a protein of several hundred residues can have a dramatic impact on its biophysical characteristics. In a broader view, these findings could have important implications for the common approach of adding tags to a protein of interest, as it is generally assumed that few additional amino acids will not have major effects on the molecule and its properties.

Material and Methods

Cell Culture

Materials and cultivation of mammalian and *Pichia* cells

All media and supplements for mammalian expression were purchased either from Sigma-Aldrich (MO, USA), Invitrogen (CA, USA) or Amimed (BioConcept, Switzerland). The antibiotics that were used, Zeocin™ and Hygromycin B, were bought from Invitrogen or PAA (Austria), respectively. All solutions used were sterilized by filtration through 0.22 µm filters (Millipore, MA, USA). Stably transformed human embryonic kidney (HEK) 293 cells were maintained in Dulbecco's Modified Eagle Medium (DMEM, Sigma-Aldrich; high glucose: 4.5 g/l) supplemented with 10% *v/v* heat-inactivated Fetal Bovine Serum (FBS, Amimed) in a humidified incubator under 5% carbon dioxide at 37°C. Expression of IgGs was carried out in DMEM supplemented with 5% *v/v* FBS (instead of the more commonly used 10%). For secretion of IgGs, derivatives of the vector pcDNA5 (Invitrogen) containing constitutively active CMV promoters upstream of the endogenous IgG signal sequences were used.

For all work with *Pichia pastoris,* the strain SMD1163 (*his4 pep4 prb1*; Invitrogen) was used. All media and supplements for this work were purchased either from Sigma-Aldrich or Invitrogen. All work was performed in a sterile laminar flow bench, and yeast growth was generally performed at 30°C. Selection of clones stably expressing the IgGs was based on Zeocin™ resistance. For secretion of IgGs, derivatives of the pGAPZαB vector (Invitrogen) containing the constitutively active *P. pastoris* GAP promoter followed by the α-factor pre-pro (αMFpp) region were used. Yeast extract-peptone-dextrose (YPD) medium containing 20 g/l peptone, 10 g/l yeast extract and 20 g/l D-glucose (plus 20 g/l agar for YPD-agar) was used for routine growth and subculturing of *Pichia* cells. Zeocin™ was added to a final concentration of 100 µg/ml. IgG expression was performed in the phosphate-buffered medium BMGY (20 g/l peptone, 10 g/l yeast extract, 100 mM potassium phosphate pH 6.0, 1.34% yeast nitrogen base (YNB) without amino acids, 1% *w/v* glycerol, 400 µg/l biotin) in the absence of antibiotic.

Construction of Expression Plasmids

Unless stated otherwise, all molecular biology methods were performed according to standard protocols.[90] All enzymes used for cloning were purchased from New England Biolabs (MA, USA) or Fermentas (Germany). In general, cloning and propagation of all plasmids was carried out in *Escherichia coli* DH5α (Life Technologies, CA, USA), grown at 37°C in low salt LB Broth (10 g/l Bacto-tryptone, 5 g/l yeast extract, 5 g/l NaCl, pH 7.5) containing 25 µg/ml Zeocin™ for the *Pichia* plasmids or in 2YT broth (16 g/l Bacto-tryptone, 10 g/l yeast extract, 5 g/l NaCl, pH 7.5) containing 100 µg/ml Ampicillin (AppliChem, Germany) for mammalian vectors.

Construction of initial expression plasmids are described in detail in Chapter 3.1. In summary, the vectors used for the creation of stable cell lines contained the genes for the light and heavy chain as complete individual expression cassettes, i.e. both genes had their own constitutively active promoter as well as their own downstream polyadenylation sequence. The attachment of additional appendices

to the N-termini of mammalian heavy and light chains was achieved by overhanging primers and PCR. The upstream regions of the heavy or light chain coding sequences were amplified with forward primers annealing upstream of the CMV promoter region and reverse primers introducing the new N-terminal appendices. The following forward primers were used for the heavy (H) or light (L) chain: HEK_H_for (5′ GCG TTT CTG GGT GAG CAA AAA CAG GAA GGC 3′) and HEK_L_for (5′ GGC ACT GTC CTC TCA TGC GTT GGG TCC 3′). The reverse primers were HEK_H-E_rev (5′ GCC AGC *CAA TTG* CAC CTG AGC TTC AGC CTC GGA CAG GAC CCA TCT GGG AGC TGC CAC CAG CAG G 3′), HEK_H-A_rev (5′ GCC AGC *CAA TTG* CAC CTG TTG AAT TCC TGC GGA CAG GAC CCA TCT GGG AGC TGC CAC CAG CAG G 3′), HEK_L-E_rev (5′ GCAT *GATATC* AGC TTC AGC CTC AGC CCA GGA TCC TGT GCC CTG AGT GAG G 3′), HEK_L-A_rev (5′ GCAT *GATATC* TTG AAT TCC TGC AGC CCA GGA TCC TGT GCC CTG AGT GAG G 3′) and HEK_L-EA_rev (5′ GCAT *GATATC* TTG AAT TCC TGC AGC TTC AGC CTC AGC CCA GGA TCC TGT GCC CTG AGT GAG G 3′). Regions coding for the new appendices are underlined while restriction sites used (*Mfe*I or *Eco*RV for the heavy or light chain, respectively) are printed in italics.

For each heavy or light chain fragment, a standard PCR using Phusion high-fidelity DNA polymerase (Finnzymes, acquired by Thermo Scientific) was carried out with a final concentration of 0.5 µM for each forward and reverse primer in the presence of 3% *v/v* DMSO. The PCR parameters were as follows: pre-denaturation at 98°C for 30 s, followed by 27 cycles of denaturation at 98°C for 10 s, annealing and extension at 72°C for 50 s, followed by a final incubation for 5 min at 72°C. The resulting PCR products were purified using a PCR clean-up kit *(*Macherey-Nagel, Germany) and digested with *Bgl*II and *Mfe*I for the heavy chain or *Afl*II and *Eco*RV for the light chain, respectively. Finally, these inserts were ligated into the identically treated and dephosphorylated initial expression vectors, yielding the final plasmids subsequently sequenced using standard techniques.

The removal of the EAEA tetrapeptide from the initial yeast heavy chain construct was performed by assembly PCR in three steps: (i) Amplification of both the region upstream of the intersection of the αMFpp with the gene encoding either the heavy or light chain as well as of the region downstream of it, both lacking the sequence encoding the EAEA tetrapeptide, (ii) assembly of fragments by an overlapping PCR method and (iii) amplification of full-length products. The upstream region was amplified using the oligonucleotides PP_out_H_for (5′ GGA AGG AGT TAG ACA ACC TGA AGT CTA GGT CCC 3′) and the reverse PP_H_rev (5′ C CAA TTG CAC CTG *TCT TTT CTC GAG AGA TAC CCC TTC TTC TTT AGC* 3′) while the downstream fragment was amplified by using PP_out_H_rev (5′ CCG GAG ACA GGG AGA GGC TCT TCT GC 3′) and PP_H_for (5′ *CT CTC GAG AAA AGA* CAG GTG CAA TTG GTA CAG TCT GGT CCG GG 3′). For the deletion of the EAEA tetrapeptide from the light chain construct, the primers were adjusted to the following: PP_out_L_for (5′ CGT GGA GGT GCA TAA TGC CAA GAC AAA GCC GC 3′), PP_L_rev (5′ GGT CAG TTC GAT ATC *TCT TTT CTC GAG AGA TAC CCC TTC TTC TTT AGC* 3′), PP_out_L_rev (5′ CTA GGA CGG TTA ACT TCG TGC CGC CGC C 3′) as well as PP_L_for (5′ *CT CTC GAG AAA AGA* GAT ATC GAA CTG ACC CAG CCG CCT TCA GTG 3′). Regions complementary to the mature N-terminus of the heavy or light chain are underlined, while those annealing to the remaining pro-region of the αMFpp are printed in italics.

The different fragments were generated by PCR using the respective forward and reverse oligonucleotides on the initial expression plasmid. The PCR was performed as follows: 30 s at 98°C; followed by 25 cycles of 10 s at 98°C and 50 s 72°C; followed by a 5 min final extension cycle at 72°C. The PCR products were purified using the same PCR purification kit as above and served as templates in the subsequent overlapping PCR. In this second PCR reaction, the outer primers (e.g. PP_out_H_for and PP_out_H_rev for the heavy chain; 0.5 µM each) were added to 15 ng of each purified PCR product. Twenty cycles of amplification (denaturation at 98°C for 10 s, annealing and extension at 72°C for 55 s) were performed, resulting in full-length fragments that were *Mlu*I/*Bst*EII or *Hpa*I/*Sac*II digested (for the heavy and light chain, respectively) and finally inserted in the appropriately cut initial vector.

The deletion of the complete pro-region in the H-pro-del construct was performed in almost the same manner, albeit using different primers in the first PCR. PP_H-pro-del_for (5′ *CC TCC GCA TTA GCT* CAG GTG CAA TTG GTA CAG TCT GGT CCG G 3′) as well as PP_H-pro-del_rev (5′ C CAA TTG CAC CTG *AGC TAA TGC GGA GGA TGC TGC GAA TAA AAC AGC* 3′) were used. The oligonucleotides for the pro-region deletion of the light chain construct were PP_L-pro-del_for (5′ *CC TCC GCA TTA GCT* GAT ATC GAA CTG ACC CAG CCG CCT TCA GTG 3′) and PP_L-pro-del_rev (5′ GGT CAG TTC GAT ATC *AGC TAA TGC GGA GGA TGC TGC GAA TAA AAC AGC* 3′). As in the primers above, regions complementary to the mature N-terminus of the heavy chain are underlined while, this time, those parts annealing to the remaining pre-region of the αMFpp are printed in italics.

Design of glycan knockout mutants T299A

To analyze the influence of the glycan moiety attached to Asn 297 in the C_H2 domain of the heavy chains, the glycan knockout construct T299A was designed. The ACG, coding for the threonine in the glycosylation motif Asn_{297}-Ser_{298}-Thr_{299}, was mutated to the alanine encoding GCG using the QuikChange® site-directed mutagenesis kit from Stratagene (acquired by Agilent, USA) according to the manufacturer's instructions. Mutagenesis was performed with the following primers encoding the mutated triplet: IgG_deglyc_for (5′ G GAG CAG TAC AAC AGC **G**CG TAC CGG GTG GTC 3′) and the complementary reverse primer IgG_deglyc_rev (5′ GAC CAC CCG GTA **C**GC GCT GTT GTA CTG CTC C 3′). After selection and sequencing, a correct clone was digested by *Eco*NI, cleaving twice in the human γ1 constant region within the C_H1 and the C_H3 domain, and the resulting fragment was inserted in the original vector cut with the same enzyme. After the selection of a construct with the correct orientation of the insert, the final vectors were re-sequenced and prepared for transformation.

Creation of stable cell lines

Flp-In HEK293 cell lines (Invitrogen), stably expressing the IgG of interest and secreting it into the culture medium, were generated according to the instructions of the Flp-In system manual. For this purpose, the expression cassettes for both the heavy and light chain genes were cloned into the pcDNA5/FRT vector and finally inserted at a specific location in the genome by the Flp recombinase within the cell. After selection, ten single clones of each construct were tested for expression of the desired IgG by Western blot analysis of the supernatant. More than 90% of the analyzed clones were positive for the desired IgG.

For the *Pichia* constructs, the corresponding plasmids containing both genes for the light and heavy chain under the control of constitutively active GAP promoters within one plasmid were linearized by *Bgl*II and integrated into the yeast genome upon transformation of competent *P. pastoris*. Finally, several colonies per construct were re-plated on fresh YPD-agar plates and subsequently analyzed for IgG expression by Western blot.

Large scale expression of IgG

For large scale expression of mammalian IgGs, 3×10^7 cells stably expressing the respective IgG were seeded in 500 cm^2 plates (Nunc), containing 65 ml of regular medium (DMEM, 10% FBS). At confluency (approximately 3 days after seeding), this medium was removed and replaced by 50 ml expression media, containing only 5% FBS. The supernatant containing the secreted IgGs was harvested 3-4 days later and renewed three times for additional rounds of expression. Detached cells were removed from the harvested supernatant by centrifugation (10 min; 9,000*g*; 4°C) and filtration through a 0.22 µm filter (Millipore). The filtered supernatant was adjusted to binding conditions by the addition of half the volume of 200 mM sodium phosphate pH 8.0 and 150 mM NaCl, refiltered and stored at 4°C until IgG purification by Protein A affinity chromatography.

Large scale expression in the yeast system was performed using a three-step procedure using two different, consecutive pre-cultures. First, 5 ml YPD medium was inoculated with a single *Pichia* colony and grown overnight. The next day, the second pre-culture of 100 ml YPD media was inoculated with 2 ml of the small scale overnight culture in a 300 ml Erlenmeyer flask. After 24 h at 30°C and 250 rpm, the final expression culture (1 l BMGY medium in 5 l Erlenmeyer flasks) was inoculated with the second pre-culture at an optical density (OD_{600}) of 1. For this inoculation, the appropriate amount of cells from the pre-culture was pelleted by centrifugation (10 min; 5,000*g*; 4°C) and resuspended in the expression media. After incubation for 48-60 h at 30°C and 250 rpm, the expression culture was centrifuged (15 min; 5,000*g*; 4°C) and the supernatant was adjusted to the conditions required for IgG purification. Since the BMGY medium already contained potassium phosphate buffer at a 100 mM concentration, no sodium phosphate but only NaCl was added to the supernatant. The pH was adjusted to 8.0 by the addition of NaOH. Afterwards, the supernatant containing the IgGs was filtered through a 0.22 µm filter and loaded onto the Protein A column.

Biophysical and biochemical analysis

Purification of IgGs

Antibodies were purified from culture supernatants by Protein A affinity chromatography. For this purpose, the supernatants - adjusted to pH 8.0 and the appropriate [Na^+] (see above) - were loaded onto HiTrap Protein A columns (GE Healthcare, USA) at 4°C with a flow rate of 1 ml/min. Chromatography was performed using an ÄKTA PrimePlus chromatography system (GE Healthcare) at 4°C. After loading, the column was washed with 100 mM sodium phosphate buffer pH 8.0 containing 150 mM NaCl. Elution of IgG was accomplished by using 0.1 M glycine pH 2.7, followed by immediate neutralization of each fraction to pH 7.5 using 1 M Tris, pH 8.0. The concentrations of the sample fractions were determined by UV-spectroscopy at 280 nm with a NanoDrop ND-1000 spectrophotometer (Thermo Scientific), assuming a mass extinction coefficient of 13.7 for a 10 mg/ml solution of IgG. The samples with the highest protein concentration were pooled and dialyzed twice against PBS (Sigma-Aldrich; 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 2.7 mM KCl, 137 mM NaCl, pH 7.1) at 4°C. After dialysis, the samples were filtered through 0.22 µm filters (Millipore) and stored at 4°C.

SDS-PAGE and Western blotting

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed using Tris-glycine 7.5% or 12% gels for non-reduced or reduced samples, respectively. Gels and samples were prepared according to standard protocols[91] and 1.5 µg of IgG antibody was loaded per lane. Non-reduced samples were only mixed with loading buffer, while for reduced samples DTT was added to a final concentration of 16 mM followed by incubation for 5 min at 95°C. As molecular mass marker the PageRuler prestained protein ladder (Fermentas) was used.

For the specific detection of the IgGs, the proteins were transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore) after their separation by SDS-PAGE. The blotting was performed following the semi-dry method within a Trans-Blot SD cell (BioRad, CA, USA) for 60 minutes at 11 V, using a Tris-glycine buffer (20 mM Tris, 150 mM glycine, 10% methanol). After blocking the membrane in 5% MPBST (5% *w/v* skimmed milk dissolved in PBST (PBS with 0.05% Tween 20) for at least 1 h, the membrane was incubated for another hour in 2.5% MPBST containing detection antibodies conjugated to alkaline phosphatase. For detection of the heavy chain, a goat anti-human Fc-specific antibody (Sigma-Aldrich, #A 9544; 1:4,000 dilution) was used, for the light chain a goat anti-human lambda-light chain-specific antibody (Sigma-Aldrich, #A 2904; 1:1,500 dilution). The polyclonal rabbit αMFpp antiserum was a kind gift of Dr. Randy Schekman (UC Berkeley) and used in a 1:1,000 dilution, followed by a secondary anti-rabbit IgG specific antibody conjugated to alkaline phosphatase (Sigma-Aldrich, #A 3687; 1:20,000 dilution). Antibodies were detected by the addition of NBT-BCIP solution (nitro-blue tetrazolium chloride with 5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt; Sigma-Aldrich) in 100 mM Tris pH 9.5, 5 mM $MgCl_2$, 100 mM NaCl; the development was terminated when the first bands became clearly visible by washing the membrane in ultra high pure water (UHP).

Isoelectric focusing (IEF)

Determination of the isoelectric point of the constructs was carried out in IEF Ready Gels (pH 3–10; Bio-Rad, CA, USA) according to the manufacturer's instructions on a Mini-PROTEAN II vertical cell system (Bio-Rad), using increasing voltages (100 V for 1 h, 250 V for 1 h, and 500 V for 30 min). Next to the samples, IEF standards, pI range 4.45 to 9.6 (Bio-Rad), were co-electrophoresed. The cathode and anode buffers were 20 mM lysine/20 mM arginine or 7 mM phosphoric acid, respectively. Following electrophoresis, the proteins were visualized by Coomassie Blue - Crocein Scarlet staining (0.04% w:v Coomassie R-250 and 0.05% w:v Crocein Scarlet 3B in 27% isopropanol / 10 % acetic acid) for 45 min and destaining (40% methanol / 10% acetic acid).

N-terminal sequence analysis

Sequence analysis was performed after *incubating 50 µg purified IgGs* with 0.3 mU of *Pfu pyroglutamate aminopeptidase* (TaKaRa Biomedicals, Japan) as described previously[92] to remove blocking pyroglutamate residues. Reduced antibodies were separated using SDS-PAGE and transferred to PVDF membranes using Tris-glycine buffer as described above. After transfer, the membrane was stained with freshly prepared Coomassie stain (0.1% Coomassie Blue R-250, 40% methanol, 1% acetic acid) for 30 s, followed by destaining with 50% methanol until protein bands were clearly visible. Protein bands corresponding to the expected MW of the heavy and light chain,

respectively, were excised and submitted for N-terminal sequencing analysis. The first 5-8 amino acids of purified IgG from *Pichia pastoris* were sequenced at the Functional Genomic Center Zurich using an Applied Biosystems Procise 492 cLC protein sequencer.

Analysis of IgG expression levels by enzyme linked immunosorbent assay (ELISA)

To analyze the influence of N-terminal variations on the IgG expression and secretion levels, the same number of HEK293 cells stably expressing the corresponding constructs was seeded in 12-well plates. Upon reaching confluency, the medium was removed and replaced by 1 ml expression media, containing 5% FBS. 24 h later, the supernatant was analyzed for its IgG content by ELISA. For the *Pichia*-produced IgGs, OD-normalized aliquots were taken after 24 h expression for ELISA analysis. For the ELISA, a capture antibody recognizing the human IgG heavy chain (Jackson ImmunoResearch, PA, USA; #209-005-098) was immobilized on MaxiSorb plates (Nunc) overnight at 4°C. After 1 h blocking in 5% skimmed milk in PBST, 100 µl of *P. pastoris*- or HEK-derived supernatant (diluted in a range of 1:25 to 1:100 in fresh BMGY or DMEM media) was incubated for 1 h at room temperature. The expressed IgG molecules were detected by incubating with an anti-human lambda-light chain specific antibody conjugated to alkaline phosphatase (Sigma-Aldrich, #A 2904; 1:2,000 dilution) for 1 h and subsequent addition of *p*-nitrophenyl phosphate (Sigma-Aldrich). Absorbance at 405 nm was measured using a Perkin Elmer HTS 7000 Plus plate reader for up to 1 h.

Circular dichroism spectroscopy

Circular dichroism (CD) measurements were performed on a Jasco J-810 Spectropolarimeter (Jasco, Japan) equipped with a computer-controlled water bath (refrigerated circulator FS18, Julabo, Germany), using a 0.5 mm cylindrical thermocuvette. CD spectra were recorded from 200 to 250 nm with a data pitch of 0.5 nm, a scan speed of 20 nm/min, a response time of 4 s, and a bandwidth of 2 nm. Measurements were performed at 25 °C and each spectrum was recorded three times and averaged. The CD signal was corrected by buffer subtraction and converted to mean residue ellipticity (MRE, θ) using the concentration of the sample determined spectrophotometrically at 280 nm. Heat denaturation curves were obtained by measuring the CD signal at 208 nm at temperatures increasing from 25°C to 90°C (heating rate 1°C/min; response time 4 s; bandwidth 2 nm). Data were collected and processed as described above. CD spectra and denaturation curves of the purified IgGs were measured in PBS (Sigma-Aldrich, pH 7.1) at a protein concentration of 5 µM.

Fluorescence spectroscopy

This method essentially was used to record the red shift of the emission maximum λ_{max} upon heat-induced unfolding of the large IgG proteins with numerous Trp residues. As water penetrates the unfolding IgG structure, the polarity in the vicinity of the Trp residues changes towards a more hydrophilic environment, thus causing an increase of λ_{max}, recorded as a loss in the F_{330}/F_{350} ratio upon unfolding. Fluorescence spectra were measured with a Jobin-Yvon Fluoromax-4 spectrofluorimeter (Horiba Scientific, NJ, USA) equipped with a Peltier-controlled cuvette holder. The temperature was controlled by an LFI3751 5A digital temperature control instrument (Wavelength Electronics Inc, MT, USA). Upon excitation at 295 nm, Trp emission spectra were recorded from 300 to 400 nm ($\Delta\lambda$ = 1 nm, scan rate 1 nm/s) in 0.5°C steps from 25°C to 90°C. The sample cuvette was equilibrated for 2 min at each temperature to ensure that the desired temperature was reached within the cell. Protein concentrations were 1 µM in every case, and all measurements were performed in PBS (pH 7.1). The intensity of the emission spectrum at 330 nm (F_{330}) and 350 nm (F_{350}) was determined at each temperature, the ratio F_{330}/F_{350} calculated and subsequently plotted as a function of temperature.

Guanidine hydrochloride-induced equilibrium unfolding

Guanidine hydrochloride (GdnHCl)-induced denaturation measurements were carried out with protein/GdnHCl mixtures containing a final protein concentration of 1 µM and denaturant concentrations ranging from 0 to 4 M (99.5% purity, Fluka MO, USA). These mixtures were prepared from a 6 M GdnHCl stock solution (in PBS, pH adjusted to 7.1) and equilibrated overnight at 25°C. Each final concentration of GdnHCl was determined by measuring the refractive index. The intrinsic fluorescence emission spectra were then recorded from 300 to 400 nm with an excitation wavelength of 295 nm. Slit widths of 2 nm were used for both excitation and emission. Individual GdnHCl blanks

were recorded and automatically subtracted from the data. The emission ratio F_{330}/F_{350} was calculated and plotted as a function of GdnHCl concentration.

Aggregation assay (light scattering at 500 nm)

The temperature of aggregation was determined by light scattering using a Jobin-Yvon Fluoromax-4 Spectrofluorimeter (Horiba Scientific). Excitation and emission wavelengths were set to 500 nm (slit width 2 nm each, integration time 1 s). Protein concentrations were 2 µM for HEK-produced IgGs and 4 µM for antibodies produced in *Pichia pastoris*. A heating rate of 1°C/min was applied starting from 25°C up to 90°C; the intensities were measured every 1°C. Each data point depicted in the plots is an average of five measured intensity values. The time period necessary to collect this set of five data points was about ~10 s, therefore only a negligible intensity change occurred due to the concomitant rise in temperature (about 0.1°C). All measurements were performed in triplicates in PBS (pH 7.1) and averaged values are given.

Deglycosylation by PNGase F treatment

The aggregation behavior was analyzed after enzymatic removal of the glycan moiety from Thr299. For this purpose, 80 µl of 2 µM purified IgGs were incubated with 50 U PNGase F (New England Biolabs, MA, USA) at 37°C for 3 h in PBS (in the absence of any detergents, in contrast to the instructions of the manufacturer). Afterwards, PNGase F-treated proteins were compared to their glycosylated counterparts by reducing SDS-PAGE for confirmation of complete deglycosylation. Then, these samples were analyzed by the aggregation assay as described before.

Size-exclusion chromatography (SEC)

SEC experiments were performed using a Superdex 200 PC 3.2/30 column (GE Healthcare). The runs were performed in PBS buffer (Sigma-Aldrich, pH 7.1) at a flow rate of 0.06 ml/min at 25°C on an ÄKTA Micro system (GE Healthcare). Samples of 50 µl containing 6.7 µM IgG were injected and protein elution was monitored at 280 nm. Cytochrome c (12.4 kDa), albumin (66 kDa) and β-amylase (200 kDa) were used as standards to calibrate the column.

Static multi-angle light scattering

The temperature-induced increase of aggregation was analyzed by SEC multi-angle light scattering (MALS). Purified IgGs were incubated in PBS (pH 7.1) at protein concentrations of 3 µM at 4°C, 37°C or 50°C for five days. Afterwards, SEC was carried out on an LC 1100 Series HPLC System (Agilent) using a Superdex 200 10/300 GL column (GE Healthcare) with a flow rate of 0.5 ml/min. MALS measurements were performed using a miniDAWN MALS instrument (Wyatt Technology, CA, USA) to determine the Rayleigh ratio. Three discrete photodetectors are spaced around the flow cell and enabled simultaneous measurements at 45°, 90° and 135°. An Optilab REX Refractometer (Wyatt Technology) was inline with the MALS detector. All measurements were carried out at 23°C. Chromatographic data were collected and processed using the ASTRA software (Version 5.3.4.14; Wyatt Technology). Bovine serum albumin (Sigma-Aldrich; 1 mg/ml) was used for the alignment and normalization of various detectors' signals relative to the 90° detector signal.

Electrospray ionization - mass spectrometry

Mass spectrometry (MS) analyses were undertaken by the Functional Genomic Center Zurich. IgG samples were reduced with DTT (50 mM final concentration), desalted using a C4 ZipTip (Millipore) and measured in 50% acetonitrile / 0.2% formic acid (pH 2). The m/z data were deconvoluted into MS-data using the MaxEnt1 software (Waters/Micromass, MA, USA).

Differential scanning calorimetry

Differential scanning calorimetry (DSC) measurements were performed using a VP-Capillary DSC system (Microcal Inc., acquired by GE Healthcare). The antibody concentrations were adjusted to 0.5 mg/ml prior to the measurement. The corresponding buffer was used as a reference. The samples were heated from 8°C to 90°C at a rate of 1°C/min after an initial 15 min of equilibration at 8°C. A filtering period of 16 s was used and data were analyzed using Origin 7.0 software (OriginLab1 corporation, MA, USA). Thermograms were corrected by subtraction of buffer-only scans and the

corrected thermograms were normalized to the molar concentration of the protein. The final excess heat capacity thermogram was obtained by interpolating a cubic baseline in the transition region. The midpoint of a thermal transition temperature (Tm) was obtained by analyzing the data using Origin 7 software provided with the instrument. As all measured transitions are irreversible, all the experimental values reported in this study for melting temperatures have to be regarded as "apparent" values.

Differential scanning fluorimetry

Differential scanning fluorimetry (DSF) was performed using the Rotor-Gene Q real-time PCR cycler (QIAGEN) and fluorescence data were collected using the instrument's HRM channel settings (λ_{ex} 460 nm; λ_{em} 510 nm). The SYPRO Orange dye (Molecular Probes) was supplied in DMSO and diluted 500-fold from the supplied stock solution into the appropriate buffers just prior to being added to the protein solutions. The samples with a final protein concentration of 3.5 µM in a 20 µl reaction mixture were subjected to a temperature ramp from 30°C to 90°C at a heating rate of 1°C/min and at 0.5°C increments with an equilibration time of 30 sec at each temperature prior to measurement. The Tm was determined as the temperature corresponding to the maximum value of the first derivative of the fluorescence changes, calculated by the software. When multiple unfolding transitions are observed, only the Tm value of the first transition can be accurately determined, as the transitions at higher temperatures overlap. Prior to the DSF analysis, several IgGs were analyzed by fluorescence spectroscopy in the presence and absence of SYPRO Orange. In agreement with the literature, the dye (at 1:200 to 1:500-fold dilutions of the original reagent) did not induce any changes in the thermal stability determined by the intrinsic fluorescence acquired during heating.

Microscale thermophoresis measurements

Binding affinities of purified IgGs to their antigen myoglobin (Sigma-Aldrich) were measured using microscale thermophoresis (MST, NanoTemper, Germany) as described previously.[93] Myoglobin was fluorescently labeled according to the manufacturer's instructions with a reactive NT-647 dye using N-hydroxysuccinimide (NHS) ester-chemistry which reacts with primary amines to form dye-protein conjugates. For each analyzed construct, a titration series with constant antigen concentration (20 µM) and varying IgG concentrations between 10^{-11} and 10^{-6} M was prepared in PBS. The mixed samples were equilibrated for 1 h at room temperature and approximately 4 µl of each sample was loaded in the capillary. An infrared laser diode within the Monolith NT.115 instrument (NanoTemper, Germany) was used to increase the temperature by 4 K in the beam center. Throughout the measurement, the fluorescence inside the capillary was recorded by a CCD-camera and the normalized fluorescence was afterwards plotted against the IgG concentrations. The K_D values were subsequently obtained from fitting the binding curves using Prism 5 (GraphPad, CA, USA).

Acknowledgements

The authors want to thank Dr. I. Jelezarov, Dr. P. Gimeson and Dr. B. Dreier for their help with the DSC and MALS analyses. We are grateful to the personnel of the Functional Genomics Center at University of Zurich for their help with MS analyses and protein sequencing. Our thanks are also due to Dr. S. Duhr from NanoTemper, Germany for his help and the chance to determine affinities by microscale thermophoresis. The polyclonal rabbit αMFpp antiserum was a kind gift of Dr. R. Schekman (UC Berkeley). We further thank Dr. Y. L. Boersma and Dr. P. Lindner for critically reading the manuscript and for valuable suggestions, and the other members of the Plückthun laboratory for fruitful discussions.

J. Schaefer was recipient of a Kekulé predoctoral fellowship of the German Chemical Industry Association and a member of the Molecular Life Science Ph.D. program. This work was supported by the Schweizerische Nationalfonds (SNF) grant 3100A0-128671/1 (to A.P.).

References

1. Plückthun, A. & Moroney, S. E. (2005). Modern antibody technology: The impact on drug development. In *Modern Biopharmaceuticals* (Knäblein, J., ed.), Vol. 3, pp. 1147-1186. 4 vols. Wiley-VCH, Weinheim.
2. Carter, P. J. (2006). Potent antibody therapeutics by design. *Nat. Rev. Immunol.* **6**, 343-357.
3. Elbakri, A., Nelson, P. N. & Abu Odeh, R. O. (2010). The state of antibody therapy. *Hum. Immunol.* **71**, 1243-1250.
4. Fontoura, P. (2010). Monoclonal antibody therapy in multiple sclerosis: Paradigm shifts and emerging challenges. *mAbs* **2**, 670-681.
5. Demarest, S. J. & Glaser, S. M. (2008). Antibody therapeutics, antibody engineering, and the merits of protein stability. *Curr. Opin. Drug Discov. Devel.* **11**, 675-687.
6. Shire, S. J. (2009). Formulation and manufacturability of biologics. *Curr. Opin. Biotechnol.* **20**, 708-714.
7. Baynes, B. M. & Trout, B. L. (2004). Rational design of solution additives for the prevention of protein aggregation. *Biophys. J.* **87**, 1631-1639.
8. Di Paolo, C., Willuda, J., Kubetzko, S., Lauffer, I., Tschudi, D., Waibel, R., Plückthun, A., Stahel, R. A. & Zangemeister-Wittke, U. (2003). A recombinant immunotoxin derived from a humanized epithelial cell adhesion molecule-specific single-chain antibody fragment has potent and selective antitumor activity. *Clin. Cancer Res.* **9**, 2837-2848.
9. Braun, A., Kwee, L., Labow, M. A. & Alsenz, J. (1997). Protein aggregates seem to play a key role among the parameters influencing the antigenicity of interferon alpha (IFN-alpha) in normal and transgenic mice. *Pharm. Res.* **14**, 1472-1478.
10. Hermeling, S., Crommelin, D. J., Schellekens, H. & Jiskoot, W. (2004). Structure-immunogenicity relationships of therapeutic proteins. *Pharm. Res.* **21**, 897-903.
11. Schellekens, H. (2005). Factors influencing the immunogenicity of therapeutic proteins. *Nephrol. Dial. Transplant.* **20 Suppl 6**, vi3-9.
12. Rosenberg, A. S. (2006). Effects of protein aggregates: an immunologic perspective. *AAPS J.* **8**, E501-507.
13. Kessler, M., Goldsmith, D. & Schellekens, H. (2006). Immunogenicity of biopharmaceuticals. *Nephrol. Dial. Transplant.* **21 Suppl 5**, v9-12.
14. Maas, C., Hermeling, S., Bouma, B., Jiskoot, W. & Gebbink, M. F. (2007). A role for protein misfolding in immunogenicity of biopharmaceuticals. *J. Biol. Chem.* **282**, 2229-2236.
15. Ring, J., Seifert, J., Jesch, F. & Brendel, W. (1977). Anaphylactoid reactions due to non-immune complex serum protein aggregates. *Monogr. Allergy* **12**, 27-35.
16. Ring, J., Stephan, W. & Brendel, W. (1979). Anaphylactoid reactions to infusions of plasma protein and human serum albumin. Role of aggregated proteins and of stabilizers added during production. *Clin. Allergy* **9**, 89-97.
17. Farid, S. S. (2006). Established bioprocesses for producing antibodies as a basis for future planning. *Adv. Biochem. Eng. Biotechnol.* **101**, 1-42.
18. Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M. & Dwek, R. A. (2007). The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu. Rev. Immunol.* **25**, 21-50.
19. Cregg, J. M., Tolstorukov, I., Kusari, A., Sunga, J., Madden, K. & Chappell, T. (2009). Expression in the yeast Pichia pastoris. *Methods Enzymol.* **463**, 169-189.
20. Cereghino, G. P., Cereghino, J. L., Ilgen, C. & Cregg, J. M. (2002). Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris. *Curr. Opin. Biotechnol.* **13**, 329-332.
21. Bretthauer, R. K. & Castellino, F. J. (1999). Glycosylation of Pichia pastoris-derived proteins. *Biotechnol. Appl. Biochem.* **30**, 193-200.
22. Montesino, R., Garcia, R., Quintero, O. & Cremata, J. A. (1998). Variation in N-linked oligosaccharide structures on heterologous proteins secreted by the methylotrophic yeast Pichia pastoris. *Protein Expr. Purif.* **14**, 197-207.
23. Jacobs, P. P., Geysens, S., Vervecken, W., Contreras, R. & Callewaert, N. (2009). Engineering complex-type N-glycosylation in Pichia pastoris using GlycoSwitch technology. *Nat. Protoc.* **4**, 58-70.
24. Vervecken, W., Callewaert, N., Kaigorodov, V., Geysens, S. & Contreras, R. (2007). Modification of the N-glycosylation pathway to produce homogeneous, human-like glycans using GlycoSwitch plasmids. *Methods Mol. Biol.* **389**, 119-138.

25. Potgieter, T. I., Cukan, M., Drummond, J. E., Houston-Cummings, N. R., Jiang, Y., Li, F., Lynaugh, H., Mallem, M., McKelvey, T. W., Mitchell, T., Nylen, A., Rittenhour, A., Stadheim, T. A., Zha, D. & d'Anjou, M. (2009). Production of monoclonal antibodies by glycoengineered Pichia pastoris. *J. Biotechnol.* **139**, 318-325.
26. Li, H., Sethuraman, N., Stadheim, T. A., Zha, D., Prinz, B., Ballew, N., Bobrowicz, P., Choi, B. K., Cook, W. J., Cukan, M., Houston-Cummings, N. R., Davidson, R., Gong, B., Hamilton, S. R., Hoopes, J. P., Jiang, Y., Kim, N., Mansfield, R., Nett, J. H., Rios, S., Strawbridge, R., Wildt, S. & Gerngross, T. U. (2006). Optimization of humanized IgGs in glycoengineered Pichia pastoris. *Nat. Biotechnol.* **24**, 210-215.
27. Cregg, J. M., Cereghino, J. L., Shi, J. & Higgins, D. R. (2000). Recombinant protein expression in Pichia pastoris. *Mol. Biotechnol.* **16**, 23-52.
28. Couderc, R. & Baratti, J. (1980). Oxidation of methanol by the yeast, Pichia pastoris. Purification and properties of the alcohol oxidase. *Agric. Biol. Chem.* **44**, 2279-2289.
29. Digan, M. E., Tschopp, J., Grinna, L., Lair, S. V., Craig, W. S., Velicelebi, G., Siegel, R., G.R., D. & Thill, G. P. (1988). Secretion of heterologous proteins from the methylotrophic yeast, Pichia pastoris. In *Development in Industrial Microbiology* (Pierce, G., ed.), Vol. 29, pp. 59-65. Elsevier Science, Amsterdam.
30. Ridder, R., Schmitz, R., Legay, F. & Gram, H. (1995). Generation of rabbit monoclonal antibody fragments from a combinatorial phage display library and their production in the yeast Pichia pastoris. *Biotechnology. (N. Y.)* **13**, 255-260.
31. Nett, J. H. (2009). Production of Antibodies in Pichia pastoris. In *Therapeutic Monoclonal Antibodies: From Bench to Clinic* (An, Z., ed.), pp. 569-584. John Wiley & Sons, Inc., Hoboken, NJ, USA.
32. Takahashi, K., Yuuki, T., Takai, T., Ra, C., Okumura, K., Yokota, T. & Okumura, Y. (2000). Production of humanized Fab fragment against human high affinity IgE receptor in Pichia pastoris. *Biosci. Biotechnol. Biochem.* **64**, 2138-2144.
33. Lange, S., Schmitt, J. & Schmid, R. D. (2001). High-yield expression of the recombinant, atrazine-specific Fab fragment K411B by the methylotrophic yeast Pichia pastoris. *J. Immunol. Methods* **255**, 103-114.
34. Gach, J. S., Maurer, M., Hahn, R., Gasser, B., Mattanovich, D., Katinger, H. & Kunert, R. (2007). High level expression of a promising anti-idiotypic antibody fragment vaccine against HIV-1 in Pichia pastoris. *J. Biotechnol.* **128**, 735-746.
35. Ogunjimi, A. A., Chandler, J. M., Gooding, C. M., Recinos, A. & Choudary, P. V. (1999). High-level secretory expression of immunologically active intact antibody from the yeast Pichia pastoris. *Biotechnol. Lett.* **21**, 561-567.
36. Ellis, S. B., Brust, P. F., Koutz, P. J., Waters, A. F., Harpold, M. M. & Gingeras, T. R. (1985). Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast Pichia pastoris. *Mol. Cell. Biol.* **5**, 1111-1121.
37. Waterham, H. R., Digan, M. E., Koutz, P. J., Lair, S. V. & Cregg, J. M. (1997). Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. *Gene* **186**, 37-44.
38. Zhang, A. L., Luo, J. X., Zhang, T. Y., Pan, Y. W., Tan, Y. H., Fu, C. Y. & Tu, F. Z. (2009). Recent advances on the GAP promoter derived expression system of Pichia pastoris. *Mol. Biol. Rep.* **36**, 1611-1619.
39. Boer, H., Teeri, T. T. & Koivula, A. (2000). Characterization of Trichoderma reesei cellobiohydrolase Cel7A secreted from Pichia pastoris using two different promoters. *Biotechnol. Bioeng.* **69**, 486-494.
40. Hohenblum, H., Gasser, B., Maurer, M., Borth, N. & Mattanovich, D. (2004). Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant Pichia pastoris. *Biotechnol. Bioeng.* **85**, 367-375.
41. Gasser, B., Maurer, M., Gach, J., Kunert, R. & Mattanovich, D. (2006). Engineering of Pichia pastoris for improved production of antibody fragments. *Biotechnol. Bioeng.* **94**, 353-361.
42. Ewert, S., Cambillau, C., Conrath, K. & Plückthun, A. (2002). Biophysical properties of camelid V_{HH} domains compared to those of human V_H3 domains. *Biochemistry* **41**, 3628-3636.
43. Knappik, A. & Plückthun, A. (1995). Engineered turns of a recombinant antibody improve its in vivo folding. *Protein Eng.* **8**, 81-89.
44. Nieba, L., Honegger, A., Krebber, C. & Plückthun, A. (1997). Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. *Protein Eng.* **10**, 435-444.
45. Bode, J., Schlake, T., Iber, M., Schubeler, D., Seibler, J., Snezhkov, E. & Nikolaev, L. (2000). The transgeneticist's toolbox: novel methods for the targeted modification of eukaryotic genomes. *Biol. Chem.* **381**, 801-813.

46. Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wölle, J., Plückthun, A. & Virnekäs, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J. Mol. Biol.* **296**, 57-86.
47. Ewert, S., Honegger, A. & Plückthun, A. (2003). Structure-based improvement of the biophysical properties of immunoglobulin VH domains with a generalizable approach. *Biochemistry* **42**, 1517-1528.
48. Jefferis, R. (2005). Glycosylation of recombinant antibody therapeutics. *Biotechnol. Prog.* **21**, 11-16.
49. Harris, R. J. (1995). Processing of C-terminal lysine and arginine residues of proteins isolated from mammalian cell culture. *J. Chromatogr. A* **705**, 129-134.
50. Kaplan, A. P., Hood, L. E., Terry, W. D. & Metzger, H. (1971). Amino terminal sequences of human immunoglobulin heavy chains. *Immunochemistry* **8**, 801-811.
51. Chelius, D., Jing, K., Lueras, A., Rehder, D. S., Dillon, T. M., Vizel, A., Rajan, R. S., Li, T., Treuheit, M. J. & Bondarenko, P. V. (2006). Formation of pyroglutamic acid from N-terminal glutamic acid in immunoglobulin gamma antibodies. *Anal. Chem.* **78**, 2370-2376.
52. Roberts, G. D., Johnson, W. P., Burman, S., Anumula, K. R. & Carr, S. A. (1995). An integrated strategy for structural characterization of the protein and carbohydrate components of monoclonal antibodies: application to anti-respiratory syncytial virus MAb. *Anal. Chem.* **67**, 3613-3625.
53. Kurjan, J. & Herskowitz, I. (1982). Structure of a yeast pheromone gene (MF alpha): a putative alpha-factor precursor contains four tandem copies of mature alpha-factor. *Cell* **30**, 933-943.
54. Kozlov, D. G. & Yagudin, T. A. (2008). Antibody fragments may be incorrectly processed in the yeast Pichia pastoris. *Biotechnol Lett* **30**, 1661-1663.
55. Emberson, L. M., Trivett, A. J., Blower, P. J. & Nicholls, P. J. (2005). Expression of an anti-CD33 single-chain antibody by Pichia pastoris. *J. Immunol. Methods* **305**, 135-151.
56. Daly, R. & Hearn, M. T. (2005). Expression of heterologous proteins in Pichia pastoris: a useful experimental tool in protein engineering and production. *J. Mol. Recognit.* **18**, 119-138.
57. Garidel, P., Hegyi, M., Bassarab, S. & Weichel, M. (2008). A rapid, sensitive and economical assessment of monoclonal antibody conformational stability by intrinsic tryptophan fluorescence spectroscopy. *Biotechnol. J.* **3**, 1201-1211.
58. Garber, E. & Demarest, S. J. (2007). A broad range of Fab stabilities within a host of therapeutic IgGs. *Biochem. Biophys. Res. Commun.* **355**, 751-757.
59. Helenius, A. & Aebi, M. (2004). Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **73**, 1019-1049.
60. Roth, J., Zuber, C., Park, S., Jang, I., Lee, Y., Kysela, K. G., Le Fourn, V., Santimaria, R., Guhl, B. & Cho, J. W. (2010). Protein N-glycosylation, protein folding, and protein quality control. *Mol. Cells* **30**, 497-506.
61. Feige, M. J., Walter, S. & Buchner, J. (2004). Folding mechanism of the CH2 antibody domain. *J. Mol. Biol.* **344**, 107-118.
62. Feige, M. J., Hendershot, L. M. & Buchner, J. (2010). How antibodies fold. *Trends Biochem. Sci.* **35**, 189-198.
63. Krapp, S., Mimura, Y., Jefferis, R., Huber, R. & Sondermann, P. (2003). Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J. Mol. Biol.* **325**, 979-989.
64. Hulett, M. D., Witort, E., Brinkworth, R. I., McKenzie, I. F. & Hogarth, P. M. (1994). Identification of the IgG binding site of the human low affinity receptor for IgG Fc gamma RII. Enhancement and ablation of binding by site-directed mutagenesis. *J. Biol. Chem.* **269**, 15287-15293.
65. Huber, R., Deisenhofer, J., Colman, P. M., Matsushima, M. & Palm, W. (1976). Crystallographic structure studies of an IgG molecule and an Fc fragment. *Nature* **264**, 415-420.
66. Sondermann, P., Huber, R., Oosthuizen, V. & Jacob, U. (2000). The 3.2-Å crystal structure of the human IgG1 Fc fragment-Fc gammaRIII complex. *Nature* **406**, 267-273.
67. Ferrara, C., Stuart, F., Sondermann, P., Brunker, P. & Umana, P. (2006). The carbohydrate at FcgammaRIIIa Asn-162. An element required for high affinity binding to non-fucosylated IgG glycoforms. *J. Biol. Chem.* **281**, 5032-5036.
68. Ravetch, J. V. & Bolland, S. (2001). IgG Fc receptors. *Annu. Rev. Immunol.* **19**, 275-290.
69. Grinna, L. S. & Tschopp, J. F. (1989). Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, Pichia pastoris. *Yeast* **5**, 107-115.

70. Martinet, W., Saelens, X., Deroo, T., Neirynck, S., Contreras, R., Min Jou, W. & Fiers, W. (1997). Protection of mice against a lethal influenza challenge by immunization with yeast-derived recombinant influenza neuraminidase. *Eur. J. Biochem.* **247**, 332-338.
71. Kayser, V., Chennamsetty, N., Voynov, V., Forrer, K., Helk, B. & Trout, B. L. (2011). Glycosylation influences on the aggregation propensity of therapeutic monoclonal antibodies. *Biotechnol. J.* **6**, 38-44.
72. Li, P., Anumanthan, A., Gao, X. G., Ilangovan, K., Suzara, V. V., Duzgunes, N. & Renugopalakrishnan, V. (2007). Expression of recombinant proteins in Pichia pastoris. *Appl. Biochem. Biotechnol.* **142**, 105-124.
73. Cereghino, J. L. & Cregg, J. M. (2000). Heterologous protein expression in the methylotrophic yeast Pichia pastoris. *FEMS Microbiol. Rev.* **24**, 45-66.
74. Brake, A. J., Merryweather, J. P., Coit, D. G., Heberlein, U. A., Masiarz, F. R., Mullenbach, G. T., Urdea, M. S., Valenzuela, P. & Barr, P. J. (1984). Alpha-factor-directed synthesis and secretion of mature foreign proteins in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 4642-4646.
75. Brocca, S., Schmidt-Dannert, C., Lotti, M., Alberghina, L. & Schmid, R. D. (1998). Design, total synthesis, and functional overexpression of the Candida rugosa lip1 gene coding for a major industrial lipase. *Protein Sci.* **7**, 1415-1422.
76. Lombardi, A., Bursomanno, S., Lopardo, T., Traini, R., Colombatti, M., Ippoliti, R., Flavell, D. J., Flavell, S. U., Ceriotti, A. & Fabbrini, M. S. (2010). Pichia pastoris as a host for secretion of toxic saporin chimeras. *FASEB J.* **24**, 253-265.
77. Monsalve, R. I., Lu, G. & King, T. P. (1999). Expressions of recombinant venom allergen, antigen 5 of yellowjacket (Vespula vulgaris) and paper wasp (Polistes annularis), in bacteria or yeast. *Protein Expr. Purif.* **16**, 410-416.
78. Brake, A. J. (1989). Secretion of heterologous proteins directed by the yeast alpha-factor leader. *Biotechnology* **13**, 269-280.
79. Piggott, J. R., Watson, M. E., Doel, S. M., Goodey, A. R. & Carter, B. L. (1987). The secretion and post translational modification of interferons from Saccharomyces cerevisiae. *Curr. Genet.* **12**, 561-567.
80. Zsebo, K. M., Lu, H. S., Fieschko, J. C., Goldstein, L., Davis, J., Duker, K., Suggs, S. V., Lai, P. H. & Bitter, G. A. (1986). Protein secretion from Saccharomyces cerevisiae directed by the prepro-alpha-factor leader region. *J. Biol. Chem.* **261**, 5858-5865.
81. Zhao, H. L., He, Q., Xue, C., Sun, B., Yao, X. Q. & Liu, Z. M. (2009). Secretory expression of glycosylated and aglycosylated mutein of onconase from Pichia pastoris using different secretion signals and their purification and characterization. *FEMS Yeast Res.* **9**, 591-599.
82. Kjeldsen, T., Brandt, J., Andersen, A. S., Egel-Mitani, M., Hach, M., Pettersson, A. F. & Vad, K. (1996). A removable spacer peptide in an alpha-factor-leader/insulin precursor fusion protein improves processing and concomitant yield of the insulin precursor in Saccharomyces cerevisiae. *Gene* **170**, 107-112.
83. Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**, 783-795.
84. Liu, Y. D., Goetze, A. M., Bass, R. B. & Flynn, G. C. (2011). N-terminal glutamate to pyroglutamate conversion in vivo for human IgG2 antibodies. *J. Biol. Chem.* **286**, 11211-11217.
85. Lawrence, M. S., Phillips, K. J. & Liu, D. R. (2007). Supercharging proteins can impart unusual resilience. *J. Am. Chem. Soc.* **129**, 10110-10112.
86. Arbabi-Ghahroudi, M., To, R., Gaudette, N., Hirama, T., Ding, W., MacKenzie, R. & Tanha, J. (2009). Aggregation-resistant VHs selected by in vitro evolution tend to have disulfide-bonded loops and acidic isoelectric points. *Protein Eng. Des. Sel.* **22**, 59-66.
87. Jespers, L., Schon, O., Famm, K. & Winter, G. (2004). Aggregation-resistant domain antibodies selected on phage by heat denaturation. *Nat. Biotechnol.* **22**, 1161-1165.
88. Perchiacca, J. M., Bhattacharya, M. & Tessier, P. M. (2011). Mutational analysis of domain antibodies reveals aggregation hotspots within and near the complementarity determining regions. *Proteins* **79**, 2637-2647.
89. Vincke, C., Loris, R., Saerens, D., Martinez-Rodriguez, S., Muyldermans, S. & Conrath, K. (2009). General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold. *J. Biol. Chem.* **284**, 3273-3284.
90. Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*. 3rd edit, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
91. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

92. Werner, W. E., Wu, S. & Mulkerrin, M. (2005). The removal of pyroglutamic acid from monoclonal antibodies without denaturation of the protein chains. *Anal. Biochem.* **342**, 120-125.
93. Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D. & Duhr, S. (2010). Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* **1**, 100.

Supplementary Material for Chapter 3.2

Charge engineering of Kex2 cleavage site

Comparing the results of the maturation of the various *Pichia* constructs, it became obvious that while the omission of the EAEA tetrapeptide did not cause any problems in the processing of the light chain (as seen in Fig. 3.2.3b bottom panel and from the sequencing results in Fig. 3.2.3c), it led to severe heterogeneity for the heavy chain H*. As the global fold of the variable domains of the light and the heavy chain including the arrangement and accessibility of the N-terminal stretch is very similar, the cause of this inefficient processing must lie elsewhere. The most obvious parameter is the primary sequence of the POI following the Kex2 cleavage site. Therefore, the sequences of the light and the heavy chain were compared with each other as well as with the "wrong" species derived from the alternative cleavage site (Fig. 3.2.3c). Comparing their amino-terminal residues, a clear difference in the charge status between the N-terminal amino-acids of the light (member of the human $V_\lambda 3$ family) and heavy chains (a human $V_H 6$ family member) were found (N-terminal sequences: QVQLV vs. DIELT). The natural human $V_\lambda 3$ sequences generally start with the sequence SYELT, while most human kappa light chains have acidic residues at their N-termini.[1] However, the consensus sequence of the human $V_\lambda 3$ family had been engineered in the HuCAL library and thus also in our case to start with DIELT for cloning reasons.[2] Interestingly, acidic amino acids are present in some V_H families as well at these positions, e.g. in $V_H 3$ and $V_H 5$.[3]

The overview depicted in Supplementary Data Fig. S3.2.2a showed that both the natural Kex2 site, the native N-terminus of the light chain as well as the alternative site within the heavy chain's pro-region exhibit acidic residues directly following the cleavage site. This suggests that the Kex2 endopeptidase seems to require negative charged residues downstream of its cutting site for efficient cleavage. This preference for acidic amino acids was also found by Bader and colleagues in their analysis of various Kex2 cleavage sites.[4] In a three-dimensional model of this enzyme, its S1' and S3' pockets are characterized by positive charges – thus allowing optimal charge-charge interactions with acidic amino acids in position P1' and P3'.

As the charge of the Kex2 substrates seems to be of great importance, the composition of the heavy chain was engineered, changing the non-charged glutamines (Q) in position P1' and P3' of the heavy chain to their charged counterparts glutamic acids (E), either individual or both at the same time. The resulting QE, EQ and EE variants were subsequently compared to the original QQ construct and analyzed for the distribution of the three species. None of these constructs led to a homogenously processed fraction but all contained molecules still having the full pro-region attached (Supplementary Data Fig. S3.2.2b). Nonetheless, the distribution of the three species could be improved by the introduction of negative charges: compared to the QQ variant the fraction of the "mature" species for the EE construct was doubled, while the percentage of the other two species (panels (c) and (d) of Supplementary Data Fig. S3.2.2) was consequently reduced.

Taken together, these data indicate that the cleavage efficiency of Kex2 is clearly influenced by the charge distribution at its substrate-binding site. Although engineering of the residues in the P1' and P3' position towards charged amino acids influenced the processing of the resulting IgG construct, it did not lead to homologous, correctly processed heavy chains. Therefore, Kex2 cleavage efficiency not only depends on the charge distribution upstream of its cleavage site, but also on other characteristics of the POI. However, we cannot yet pinpoint what these features are exactly.

References

1. Kawasaki, K., Minoshima, S., Nakato, E., Shibuya, K., Shintani, A., Schmeits, J. L., Wang, J. & Shimizu, N. (1997). One-megabase sequence analysis of the human immunoglobulin lambda gene locus. *Genome Res.* **7**, 250-261.
2. Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wölle, J., Plückthun, A. & Virnekäs, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J. Mol. Biol.* **296**, 57-86.
3. Matsuda, F., Ishii, K., Bourvagnet, P., Kuma, K., Hayashida, H., Miyata, T. & Honjo, T. (1998). The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. *J. Exp. Med.* **188**, 2151-2162.
4. Bader, O., Krauke, Y. & Hube, B. (2008). Processing of predicted substrates of fungal Kex2 proteinases from Candida albicans, C. glabrata, Saccharomyces cerevisiae and Pichia pastoris. *BMC Microbiol.* **8**, 116.

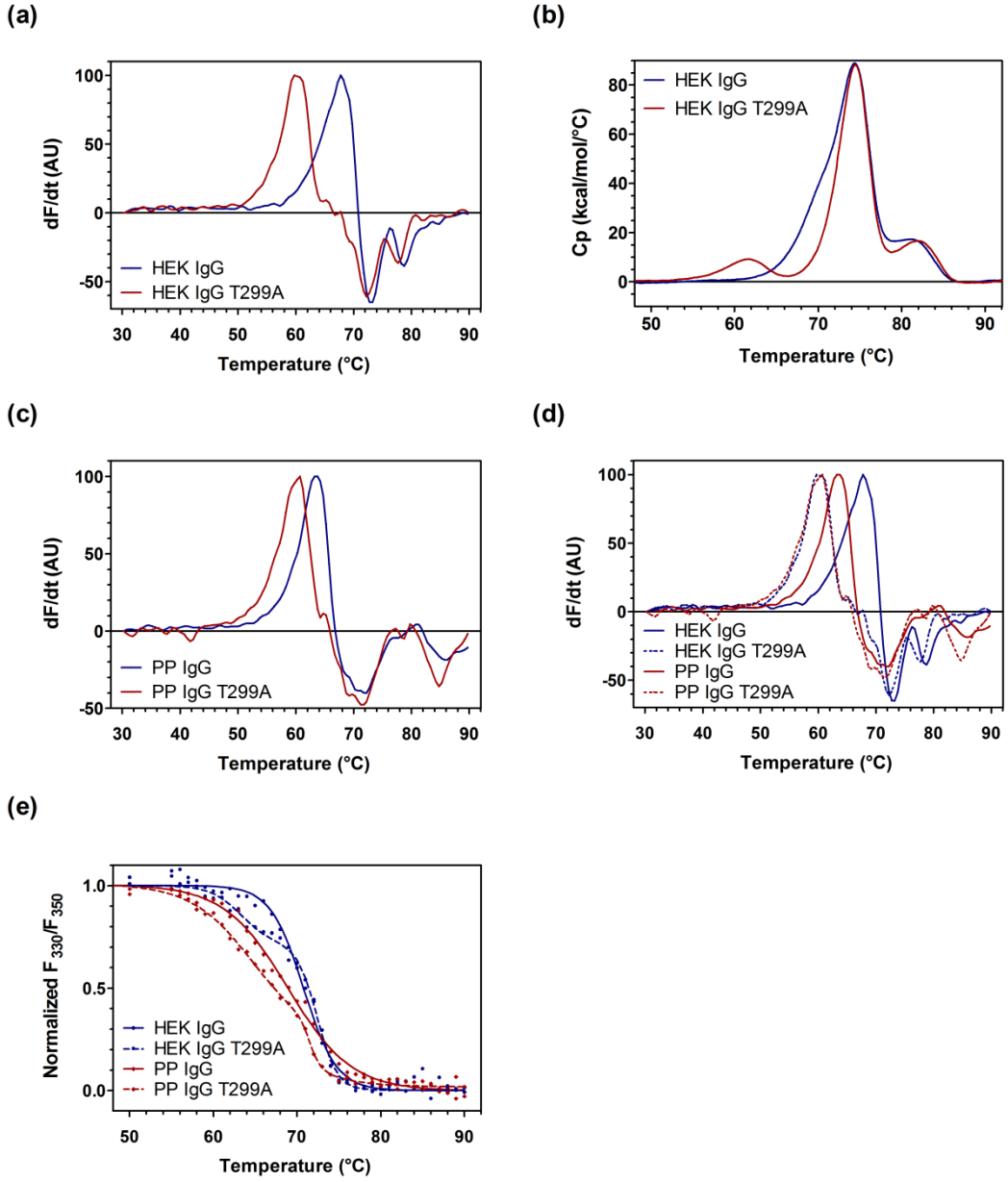


Fig. S3.2.1. Influence of glycosylation on IgG stability. (a) Comparison of the DSF signal of HEK-produced native IgG with that of its unglycosylated counterpart (T299A glycan knock-out). The transition of the C_H2 unfolding is the dominant signal detected. **(b)** DSC signals recorded for the same constructs as in (a). **(c)** DSF analysis of *Pichia*-produced IgG constructs. **(d)** Comparison of DSF signals recorded for HEK- and *Pichia*-produced IgGs in their glycosylated and un-glycosylated state. **(e)** Thermal denaturation curves of the constructs seen in (d) measured by intrinsic tryptophan fluorescence. The curves were derived from the intensity ratio of the emission spectrum at 330 nm and 350 nm upon excitation at 295 nm, normalized to the plateaus of fully folded and fully denatured proteins. This value is plotted as a function of temperature.

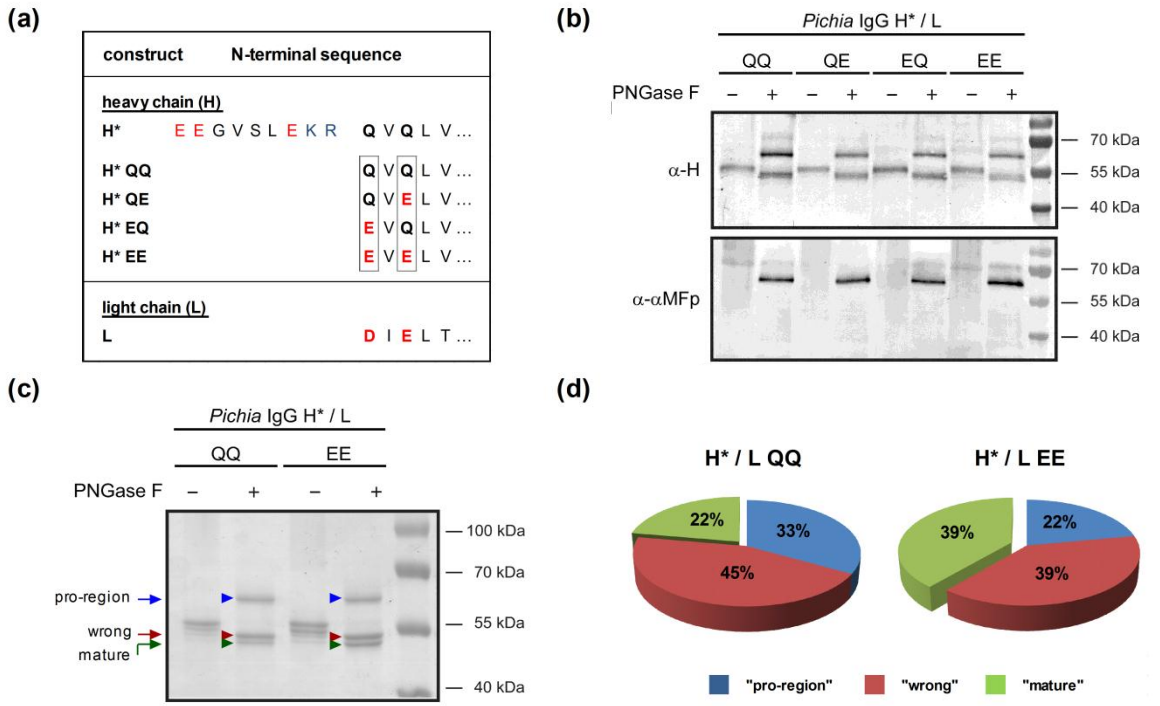


Fig. S3.2.2. Charge engineering of *Pichia*-produced IgG H* / L construct. (a) Comparison of the mature N-terminal amino acid sequences of different yeast heavy chain variants analyzed. The uppermost sequence is the one primarily found for the H* construct, derived from an alternative cleavage within the pro-region and carrying two negatively charged amino acids at its N-terminus (red color). To test charge influence on cleavage efficiency the residues at the first and third position of the mature heavy chain N-terminus were mutated from glutamine (Q) to glutamic acid (E) in individual combinations or combined in the EE variant. **(b)** Western blot analysis of IgG H* / L charge variants after incubation in the absence or presence of PNGase F. Shown are the blots detected with antibodies specific to the heavy chain (α-H) and to the α-factor pre-pro-region (α-αMFpp). **(c)** Comparative reducing SDS-PAGE analysis of yeast IgG H* / L QQ and EE variants after incubation in the absence or presence of PNGase F, stained with Coomassie Blue. **(d)** Comparison of distribution of IgG H* / L forms for the QQ and EE variants, based on the relative signal intensities found in Edman sequencing.

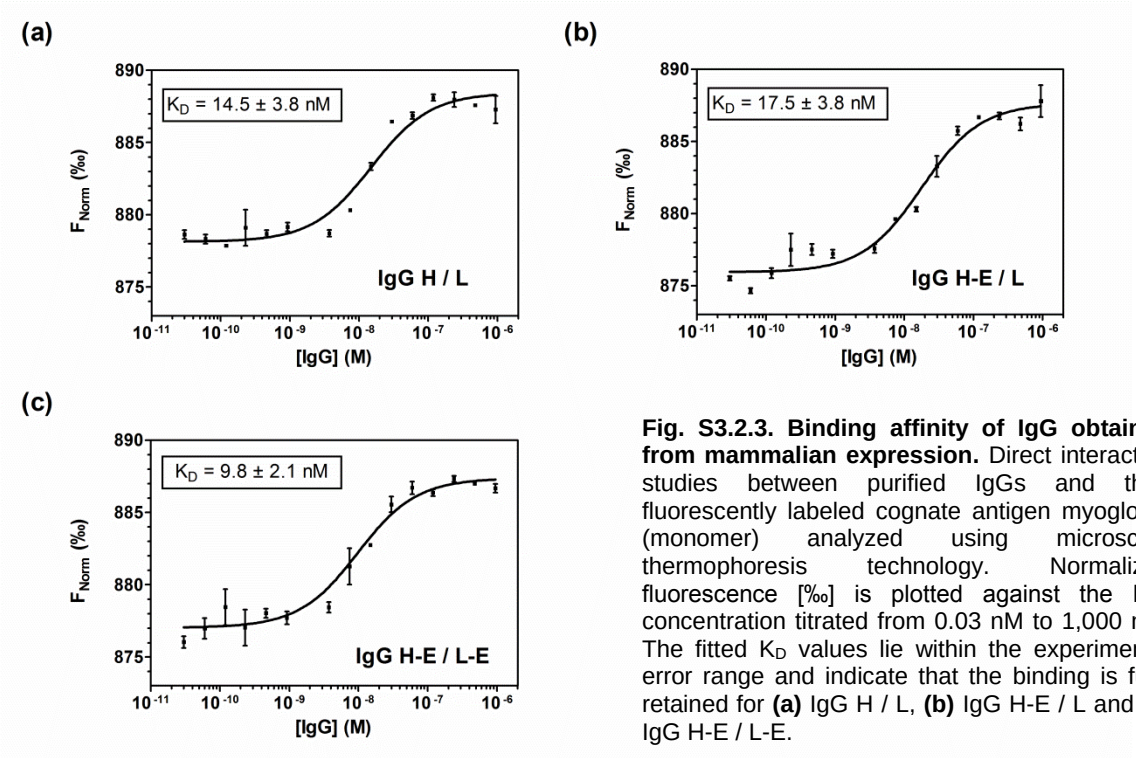


Fig. S3.2.3. Binding affinity of IgG obtained from mammalian expression. Direct interaction studies between purified IgGs and their fluorescently labeled cognate antigen myoglobin (monomer) analyzed using microscale thermophoresis technology. Normalized fluorescence [‰] is plotted against the IgG concentration titrated from 0.03 nM to 1,000 nM. The fitted K_D values lie within the experimental error range and indicate that the binding is fully retained for **(a)** IgG H / L, **(b)** IgG H-E / L and **(c)** IgG H-E / L-E.

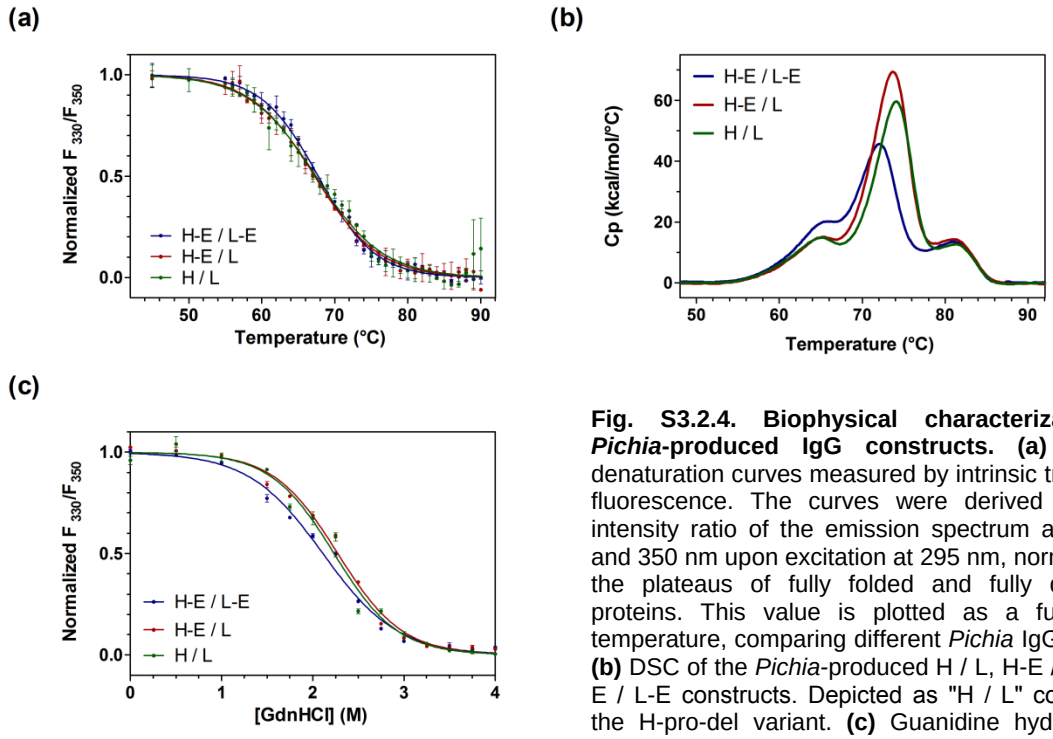


Fig. S3.2.4. Biophysical characterization of *Pichia*-produced IgG constructs. (a) Thermal denaturation curves measured by intrinsic tryptophan fluorescence. The curves were derived from the intensity ratio of the emission spectrum at 330 nm and 350 nm upon excitation at 295 nm, normalized to the plateaus of fully folded and fully denatured proteins. This value is plotted as a function of temperature, comparing different *Pichia* IgG variants. **(b)** DSC of the *Pichia*-produced H / L, H-E / L and H-E / L-E constructs. Depicted as "H / L" construct is the H-pro-del variant. **(c)** Guanidine hydrochloride (GdnHCl)-induced denaturation of IgGs expressed in *Pichia pastoris*. The denaturation was followed by plotting the F_{330}/F_{350} ratio normalized to the plateaus of fully folded and fully denatured proteins as a function of GdnHCl concentration.

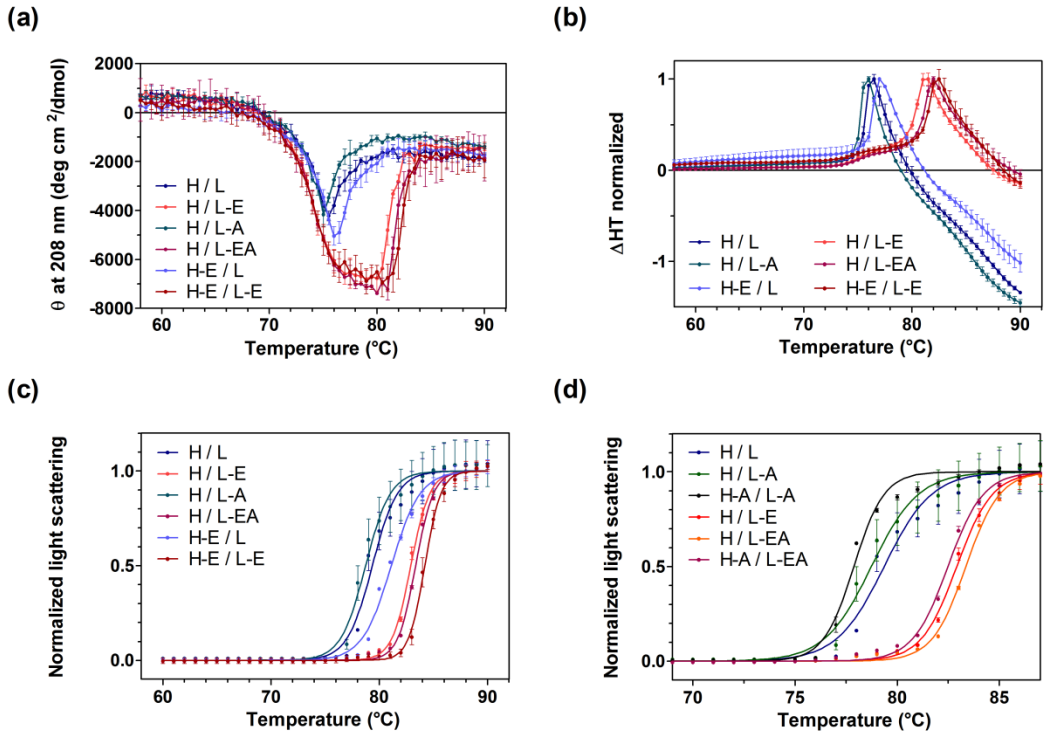


Fig. S3.2.5. Aggregation behavior of different IgGs obtained from mammalian expression. (a) Thermal denaturation curves. The denaturation was followed by CD, plotting the signals at 208 nm as a function of temperature. A comparison between different mammalian IgG variants is shown and the values are reported as MRE. **(b)** High tension (HT) voltage signal corresponding to the curves shown in (a) recorded by the photomultiplier connected to the CD spectrometer. The values were normalized by setting the initial and the highest values of the curves to 0 and 1, respectively. **(c)** Aggregation of IgG constructs

measured by light scattering at 500 nm. Emission was recorded at the same wavelength as excitation. **(d)** Zoom at the aggregation temperatures of selected constructs. In addition to some of the constructs already shown in (c), also the signals derived from IgGs carrying the control peptide AGIQ on the heavy chain (H-A format) are presented.

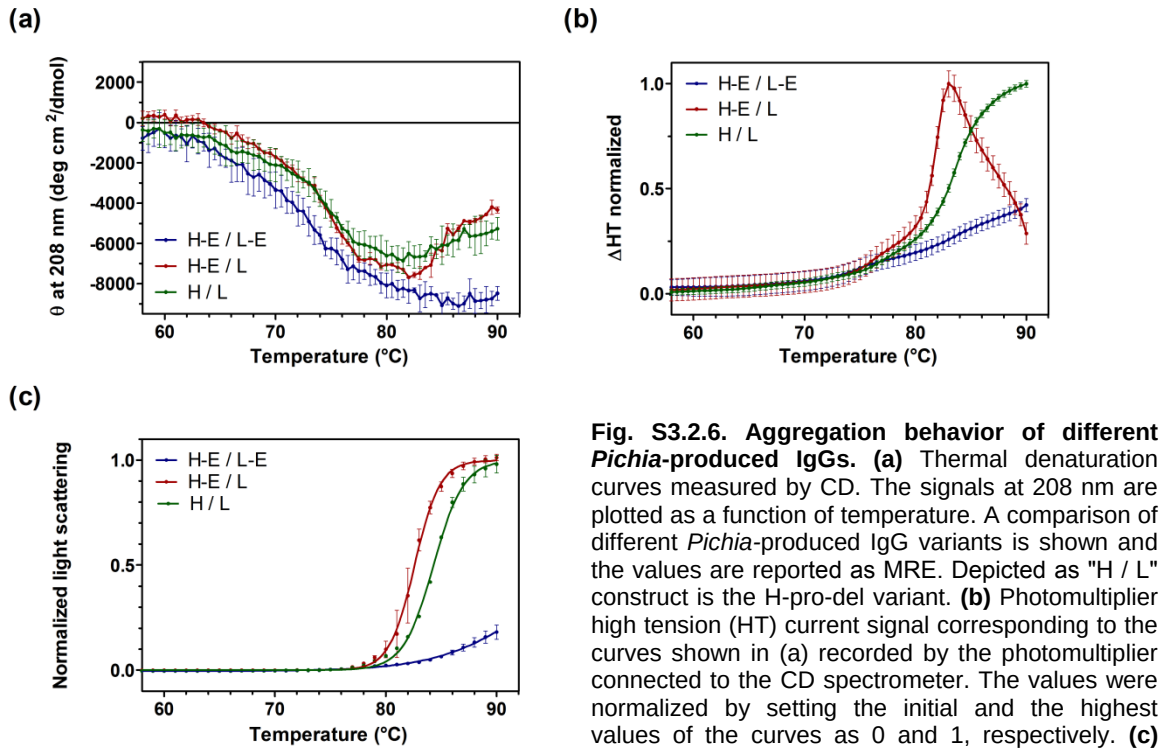


Fig. S3.2.6. Aggregation behavior of different *Pichia*-produced IgGs. (a) Thermal denaturation curves measured by CD. The signals at 208 nm are plotted as a function of temperature. A comparison of different *Pichia*-produced IgG variants is shown and the values are reported as MRE. Depicted as "H / L" construct is the H-pro-del variant. **(b)** Photomultiplier high tension (HT) current signal corresponding to the curves shown in (a) recorded by the photomultiplier connected to the CD spectrometer. The values were normalized by setting the initial and the highest values of the curves as 0 and 1, respectively. **(c)** Aggregation of IgG constructs measured by light scattering at 500 nm.

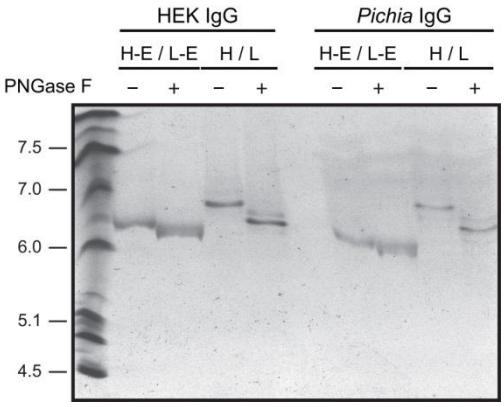


Fig. S3.2.7. Isoelectric Focusing of HEK and *Pichia* produced IgGs Non-reduced IgGs were separated based on their isoelectric point (pI) after incubation in the absence or presence of PNGase F. A pI standard is shown in the left lane, with the corresponding pI values denoted next to it. Based on the primary sequences, the pI's were calculated to be 6.1 for the H-E / L-E and 6.7 for the H / L constructs, respectively. Upon deglycosylation, the asparagine to which the glycan was attached is converted to aspartic acid, resulting in an additional negative charge to the protein, which consequently focuses at a slightly lower pI value.

3.3 Construction of scFv Fragments from Hybridoma or Spleen Cells by PCR Assembly

Jonas V. Schaefer, Annemarie Honegger and Andreas Plückthun

in **Antibody Engineering** (Kontermann, R., and Dübel, S., eds) Vol. 1, 2nd edit., 2010, pp. 21-44, Springer-Verlag, Berlin Heidelberg, Germany

Chapter 3
Construction of scFv Fragments from Hybridoma or Spleen Cells by PCR Assembly

Jonas V. Schaefer, Annemarie Honegger, and Andreas Plückthun

Abbreviations

BSA	Bovine serum albumin
DMSO	Dimethylsulfoxide
HRP	Horse radish peroxidase
IPTG	Isopropylthiogalactoside
PEG	Polyethylene glycol
PBS	Phosphate buffered saline
scFv	Single-chain Fv fragment
cfu	Colony forming units
tet	Tetracycline

3.1 Introduction

Today, antibodies can be obtained from naive repertoires (Winter et al. 1994; Vaughan et al. 1996) or libraries of fully synthetic genes (Knappik et al. 2000), and in the last decade, numerous libraries have been described (reviewed in Mondon et al. 2008). Nonetheless, hybridomas have remained the predominant source of antibodies, and a wealth of well characterized and even unique clones exist and are continuing to be generated. There is, thus, great interest in immortalizing these clones, in the extreme case, as a computer file of the sequences, as well as in accessing the antibody in a variety of new formats. To obtain enough material

J.V. Schaefer, A. Honegger, and A. Plückthun (✉)
Biochemisches Institut, Universität Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland
e-mail: plueckthun@bioc.uzh.ch

for detailed biochemical and biophysical analyses of the deduced antibodies after immunization, their cloning into formats compatible with recombinant expression is beneficial, if not essential. For this purpose, the antibody genes must be cloned, and the binding properties of the recombinant protein have to be verified. In addition to existing hybridomas, the immune response of an animal upon exposure to various antigens may often be of particular scientific interest in itself and also lead to the discovery of new and potent binders. Therefore, there is merit in immortalizing the results from new immunizations as well. In this case, it is not necessary to take the detour of first making hybridomas, but instead, mRNA isolated from spleen can be directly used for the creation of an immune library, from which binders can be subsequently isolated by phage display and their sequences determined.

Once the antibody genes have been successfully cloned and after the presumed binding properties of the recombinant antibodies have been experimentally verified, their sequences can be used for modeling (www.bioc.uzh.ch/antibody/), and their structure subsequently be determined by crystallography (Honegger et al. 2005) or NMR (Freund et al. 1994; Tugarinov et al. 2000). The recombinant single-chain Fv format (Huston et al. 1988; Glockshuber et al. 1990) is an ideal starting point for all engineering efforts, from sensors (Backmann et al. 2005; Morfill et al. 2007) to therapeutic fusion proteins (Di Paolo et al. 2003), or imaging reagents (Adams et al. 1993) to multivalent and multispecific reagents (Plückthun and Pack 1997), just to name a few illustrative examples. Recombinant expression of these proteins also allows one to evolve the affinity further than the immune system normally does, e.g., to low picomolar K_D for scFv fragments (Zahnd et al. 2004; Luginbühl et al. 2006). Finally, some natural antibodies may not be of sufficient stability, which can also be corrected by engineering (Wörn and Plückthun 2001; Ewert et al. 2004). In addition, the murine antibody can be humanized for its use in therapy – a procedure rapidly achievable at the scFv stage.

The key prerequisite for the use of recombinant antibody technologies, starting from immune repertoires or defined hybridomas, is the reliable cloning of functional immunoglobulin genes. Even though hybridomas are considered to express "monoclonal" antibodies, hybridoma clones may encode more than one functional or even nonfunctional heavy or light chains (Kütemeier et al. 1992). As has been reported previously, several kappa chain-secreting hybridomas, possessing X63Ag8.653 myeloma cells as fusion partner, also occasionally transcribe a functional lambda chain, competing with the V_κ gene for in-frame scFv antibody assembly (Krebber et al. 1997). As these false or heterogeneous genes might also be amplified and subsequently assembled into the scFv fragments, it is highly recommended to include an enrichment procedure in the cloning protocol. This step can be circumvented and replaced by screening of clones at the scFv level, but the phage enrichment is generally much faster if incorrect sequences abound. Obviously, selection by phage display or by another selection technology such as ribosome display (Hanes and Plückthun 1997; Hanes et al. 1998; for detailed protocols see Schaffitzel et al. 2005; Amstutz et al. 2006; Zahnd et al. 2007) is mandatory when starting from spleens of immunized mice.

This chapter largely follows our earlier protocols (Plückthun et al. 1996; Krebber et al. 1997; Burmester and Plückthun 2001). A number of variable antibody domains of hybridomas were accessible with those procedures and reagents whose genes could not be cloned in other experimental setups. The present protocol is based on a standard phage display system, which was optimized for robustness, vector stability, and directional cloning using a single rare cutting restriction enzyme as well as tight control of the expression of the scFv-gene III fusion (Krebber et al. 1997). As the procedures for the construction of scFv fragment libraries from immunized mice and that of cloning one specific antibody from hybridomas are essentially the same, we combined them in just one protocol. However, there are slight differences in the initial preparation of the cells, and high ligation and transformation yields for library cloning are, of course, essential, as explained under "notes."

The current version of this protocol contains improvements in the methods but, most importantly, newly designed primer sequences for the amplification of V_H and V_L genes. They are based on our analysis of a reference set of murine germline sequences found in the most recent version of the IMGT database (http://imgt.cines.fr/textes/vquest/refseqh.html), which thus incorporates most of the knowledge of the mouse genome (for a description of the original database, see Lefranc and Lefranc 2001). Our key criterion was a faithful amplification of the variable region genes preserving as much sequence identity as possible, avoiding the generation of nonnatural residue combinations, which could result in sequences problematic for folding and stability (Honegger and Plückthun 2001; Jung et al. 2001). We also tried to ensure similar annealing temperatures with the different genes, as well as keeping the degeneracy on the DNA level as small as possible. Furthermore, we avoided pronounced secondary structures within the oligonucleotides such as hairpin loops or primer-dimers (which were checked against themselves using the appropriate analysis tools in the Vector NTI software (Invitrogen)). The primers shown below are the result of this iterative process and have also been tested with a slightly different overhang.

The cloning strategy outlined in this protocol (Fig. 3.1) allows the simple conversion of the expression format from the initial scFv fragments to other formats and fusion proteins. Insertion of the assembled scFv gene into the described standard vectors pAK100 and pJB12 leads to the expression of a scFv-gene III fusion applicable for phage display, due to read-through of the amber codons whenever expressed in strains with amber suppressor tRNA such as *Escherichia coli* XL1-Blue. In bacterial strains lacking such suppressor tRNA, the amber stop codons result in translation termination and production of unfused scFv fragments. For purposes of IMAC purification or whenever other fusions will be constructed, it is, however, advantageous to reclone the fragments directly into appropriate vectors (Figs. 3.4 and 3.5) (Plückthun et al. 1996), carrying stronger translation initiation sites. Conversely, it is not advantageous to make expression too strong for phage display, as discussed below. Although not explicitly mentioned, a very similar strategy of cloning (Fig. 3.1) only requiring altered reverse primers can be used for the design of Fab versions of the desired antibodies.

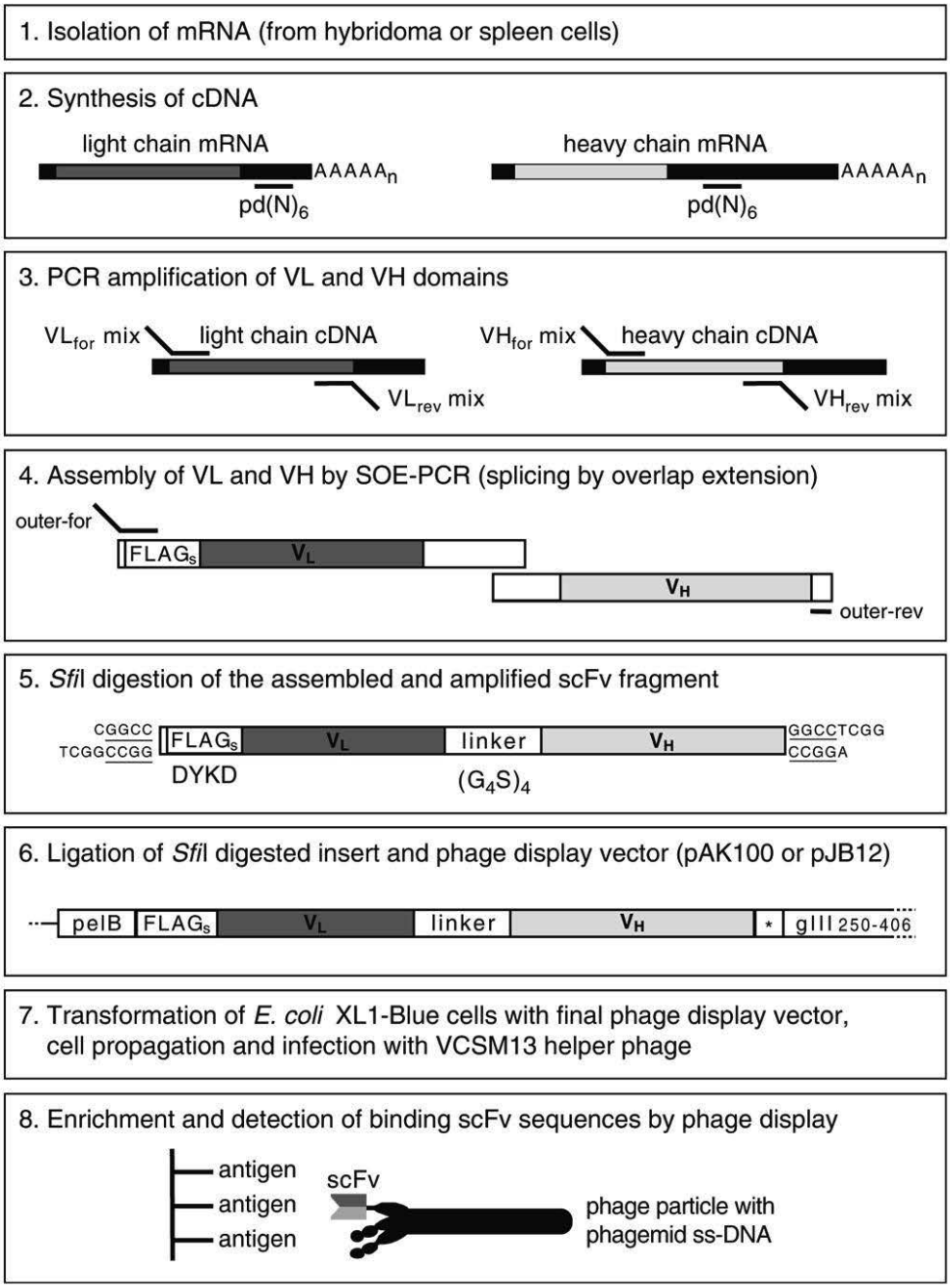


Fig. 3.1 *Schematic overview of the amplification and cloning procedure*. After its isolation from hybridoma or spleen cells, the mRNA provides the basis for cDNA synthesis, utilizing random hexamer primers. The cDNA is used afterward as template for PCR amplification of V_L and V_H domains (symbolized by the gray boxes, not drawn to scale), which are subsequently assembled by SOE-PCR into the scFv format by the outer primer pair outer-for and outer-rev. For antibody cloning into the phagemid, only the rare cutting enzyme *Sfi*I is used, guaranteeing directional cloning due to the resulting different overhangs at the cleavage site as indicated. In addition, self-ligation of insert or vector molecules is excluded by the asymmetry generated in the cut vector. $FLAG_s$ indicates the shortened *N*-terminal 4-amino acid FLAG tag (Knappik and Plückthun 1994)

3.2 Materials

- 5×10^6 cells from a growing or frozen hybridoma culture or spleen cells, respectively
- PCR primers (Figs. 3.2 and 3.3) and corresponding plasmids (Figs. 3.4 and 3.5)
- Helper phage (e.g., Stratagene VCSM13 # 200251)
- F^+, *sup*E, *rec*A strain (e.g., *E. coli* XL1-Blue) (available in electrocompetent/chemocompetent form from Stratagene)
- Anti-M13 antibody HRP-conjugate (GE Healthcare; # 27-9421-01)
- PEG 6000 (Fluka)
- Sterile, RNase-free equipment: pipet tips, tubes, RNase-free ultra high purity (UHP) water, baked nondisposable glassware, and sterile, disposable plasticware
- Standard molecular biology equipment and reagents for:
  - Determining the isotype of mAbs (Roche IsoStrip Mouse Monoclonal Antibody Isotyping Kit)
  - Purifying RNA (Invitrogen TRIzol reagent and Qiagen RNeasy Mini Kit)
  - Performing a cDNA synthesis reaction (Qiagen QuantiTect Reverse Transcription Kit)
  - Performing PCR reactions
  - Purifying PCR products (Macherey Nagel PCR clean-up Gel Extraction Kit)
  - Cutting and gel-purifying DNA (Sigma-Aldrich GenElute Gel Extraction Kit)
  - Concentrating DNA (Amicon Microcon 30 for volumes less than 500 µl)
  - Ligating and transforming DNA
  - Growing bacteria and phages
  - Conducting an Enzyme Linked Immunosorbent Assay (ELISA)
  - Performing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and subsequent immunoblotting.

Fig. 3.1 (continued) and the asterisk symbolizes either the myc tag or the trypsin cleavage site, present in pAK100 and pJB12, respectively. After infection with VCSM13 helper phages, the transformed XL1-Blue cells produce phages, displaying the scFv antibody on their surface. The subsequent enrichment of these phages by panning against the antigen allows the selection of functional antibody sequences from the library generated from the spleen cells. In addition, this approach also supports the isolation of specific scFv fragments if the hybridoma cell line initially contained only a small fraction of mRNA coding for this particular antibody. Phage ELISA then identifies the antigen-binding clones. Subsequently, the binding properties of the unfused scFv in the absence of phage need to be verified after recloning into a more powerful expression vector (Figs. 3.4 and 3.5) and purification from *E. coli* (not shown in the diagram)

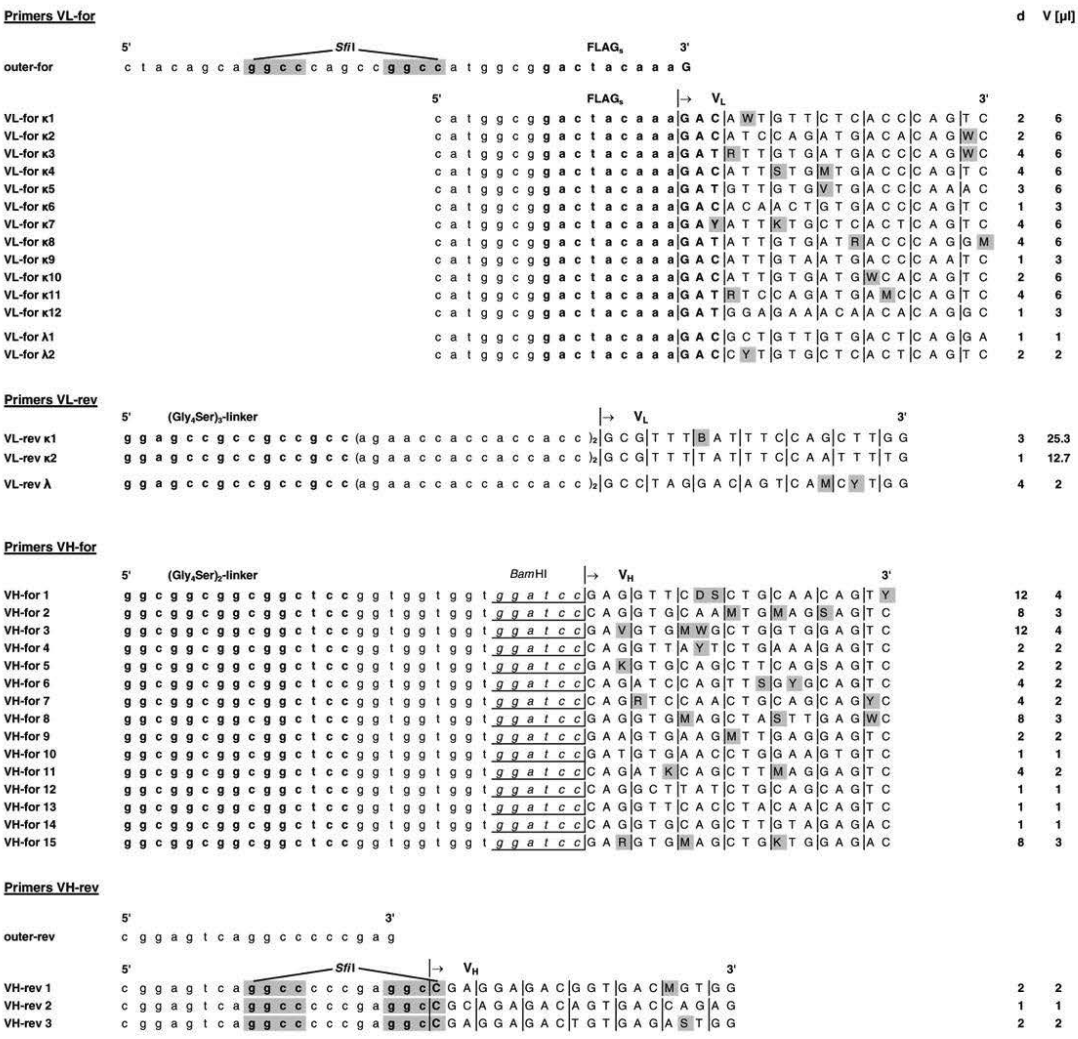


Fig. 3.2 *List of primers used for assembling mouse scFv fragments*. The depicted oligonucleotides direct the assembly of scFv fragments in the orientation V_L-$(G_4S)_4$-V_H, being compatible with the vectors presented in Fig. 3.4. They have been newly derived from an analysis of the complete set of mouse sequences, and are thus different from the previously reported sets (Burmester and Plückthun 2001; Krebber et al. 1997). The sequences are given using the IUPAC nomenclature of mixed bases (shown as capital letters with gray background, R = A or G; Y = C or T; M = A or C; K = G or T; S = C or G; W = A or T; H = A or C or T; B = C or G or T; V = A or C or G; D = A or G or T). A column lists the d-fold degeneration encoded in each primer, d being the number of unique sequences. The "VL-for" primers VL-for κ1 to VL-for κ12 encode a stretch of 20 bases, hybridizing to the mature mouse antibody κ sequences (in capital letters). The preceding sequence that encodes the shortened FLAG sequence (Knappik and Plückthun 1994) is shown in bold. Since the FLAG tag uses the fixed *N*-terminal aspartate of the mature antibody (encoded by GAY), only three additional amino acids are necessary. The FLAG codons are then preceded by the codons specifying the end of the *pelB* signal sequence. The "VL-for" primers VL-for λ1 and VL-for λ2 for mouse lambda chains are constructed in an analogous manner (the *N*-terminal glutamate of the mature mouse λ sequence is replaced by aspartate (encoded by GAC) to generate a FLAG tag). The "VL-rev" primer sequences are complementary to the J-elements of kappa or lambda chains (capital letters) and encode three repeats of the Gly_4Ser sequence, with the terminal one (bold) possessing a different codon usage to minimize incorrect overlaps during the PCR

3.3 Method

3.3.1 Isolation of mRNA and cDNA Synthesis

1. Take 5×10^6 cells from a frozen or growing hybridoma culture (for isotype determination, use the Roche IsoStrip Mouse Monoclonal Antibody Isotyping Kit) or spleen cells, respectively (see note). Perform a total RNA preparation, combining homogenization of cells in the presence of TRIzol Reagent (Invitrogen) with RNA purification using the Qiagen RNeasy Mini Kit as described by the manufacturers. According to the supplier, the latter kit can be used for up to 1×10^7 cells, but in order to get highly pure mRNA, take only 5×10^6 cells.

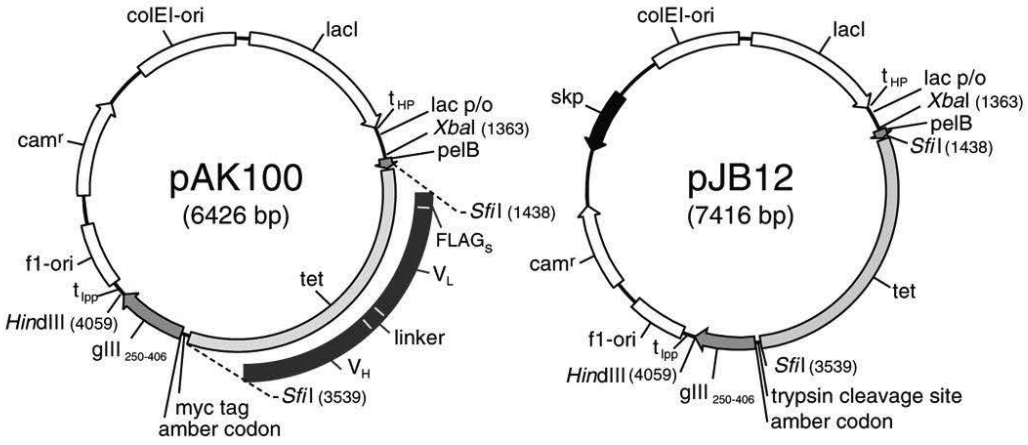
   Note: For RNA preparation from mouse spleen (typically yielding 5×10^7 B-cells each), first separate it from connective tissue with sharp forceps or scissors (if frozen, also cut the frozen tissue into smaller pieces and pulverize using a mortar) and homogenize it using the Tissue Lyser (Qiagen) or similar homogenizers in the presence of 1 ml TRIzol Reagent per 50 mg of tissue. Make sure not to use too many cells as spleens are typically rich in nucleases, and, therefore, enough RNAse-deactivating components from the TRIzol Reagent should be present in the solution. TRIzol Reagent is a commercial monophasic preparation of guanidinium isothiocyanate and phenol and only the addition of chloroform separates the solution in two phases. If desired, polyA$^+$ mRNA can subsequently be isolated from the total RNA using the Oligotex Direct mRNA Mini Kit (Qiagen) – however, in most cases, this should not be necessary for the subsequent production of cDNA. Therefore, we do not recommend including this additional purification step, as it might lead to loss of mRNAs present only in smaller quantities. Since specific amplification primers are used, we consider it rather advantageous to work with a higher total amount of RNA.
2. Separate the RNA from DNA and proteins by phenol-chloroform extraction with subsequent silica membrane purification as described by the manufacturer (Invitrogen). Transfer the upper aqueous phase to a new, RNase-free tube. Add an equal volume of 100% ethanol dropwise, as its presence is required for the RNeasy columns to bind the RNA during the initial application. Transfer up to 700 µl of the mixture, including any precipitate that may have formed,

Fig. 3.2 (continued) assembly reaction. Please note that for these primers, the two identical linker repeats are presented by a parenthesis with the subscript 2. The "VH-for" primers encode the other part of the linker (overlap with VL-rev shown in *bold*) as well as a *Bam*HI recognition site (*underlined*). The 20 bases given in capital letters hybridize with the mature mouse VH sequences. The last 20 nucleotides (nt) at the 3′ end of the "VH-rev" primers hybridize with the JH region. The first nt shown in capital letters will cause a silent mutation at the end of V_H in order to code for the first nt of the second *Sfi*I recognition site (*bold and highlighted*). The final assembly of the scFv gene by SOE-PCR is carried out with the outer-for and outer-rev primer set. The outer primer outer-for encodes the first *Sfi*I site (*bold and highlighted*). The last column lists the volume that should be used when mixing the primers (see text)

V_L	1	2	3	4	5	6	7	amino acid position	102	103	104	105	106	107	108	V_L
VL-for κ1	D	I,N	V	L	T	Q	S		T	K	L	E	I,M	K	R	VL-rev κ1
VL-for κ2	D	I	Q	M	T	Q	S,T		T	K	L	E	I	K	R	VL-rev κ2
VL-for κ3	D	I,V	V	M	T	Q	S,T		A,T	K,R	L,V	T	V	L	G	VL-rev λ
VL-for κ4	D	I	L,V	L,M	T	Q	S									
VL-for κ5	D	V	V	L,M,V	T	Q	T									
VL-for κ6	D	T	T	V	T	Q	S									
VL-for κ7	D	I	L,V	L	T	Q	S									
VL-for κ8	D	I	V	I,M	T	Q	A,D									
VL-for κ9	D	I	V	M	T	Q	S									
VL-for κ10	D	I	V	M	S,T	Q	S									
VL-for κ11	D	I,V	Q	M	N,T	Q	S									
VL-for κ12	D	G	E	T	T	Q	A									
VL-for λ1	D	A	V	V	T	Q	E									
VL-for λ2	D	P,V	V	L	T	Q	S									

V_H	1	2	3	4	5	6	7	amino acid position	107	108	109	110	111	112	113	V_H
VH-for 1	E	V	H,L,Q,R	L	Q	Q	F,L,S		T	T	V	T	V	S	S	VH-rev 1
VH-for 2	Q	V	Q	L,M	K,Q	E,Q	S		T	L	V	T	V	S	A	VH-rev 2
VH-for 3	D,E	V	K,M,Q	L	V	E	S		T	S,T	L	T	V	S	S	VH-rev 3
VH-for 4	Q	V	I,T	L	K	E	S									
VH-for 5	D,E	V	Q	L	Q	E,Q	S									
VH-for 6	Q	I	Q	F,L	A,V	Q	S									
VH-for 7	Q	I,V	Q	L	Q	Q	P,S									
VH-for 8	E	V	K,Q	L	L,V	E	S,T									
VH-for 9	E	V	K	I,L	E	E	S									
VH-for 10	D	V	N	L	E	V	S									
VH-for 11	Q	I,M	Q	L	K,Q	E	S									
VH-for 12	Q	A	Y	L	Q	Q	S									
VH-for 13	Q	V	H	L	Q	Q	S									
VH-for 14	Q	V	Q	L	V	E	T									
VH-for 15	E	V	K,Q	L	V,L	E	T									

Fig. 3.3 *Deduced amino acid sequence of the V domain, hybridizing part of primers from Fig. 3.2*. The residues of V_L and V_H are numbered according to Kabat et al. (1991). All forward primers determine only the first 2 nt of residue 7, whereas the reverse primers determine the last 2 nt of position 102 or 107, respectively. Therefore, the alternative translations at those latter positions given do not indicate that they encode a mixed codon



pAK / pJB vector series		phage display	Skp coexpression	enhanced expression	trypsin cleavage site	IMAC purification	C-terminal detection	direct detection	dimerization
pAK100	XbaI SfiI SfiI EcoRI HindIII — pelB tet myc * gIII 250-406	●							
pAK300	XbaI SfiI SfiI HindIII — pelB tet (His)6					●	●		
pAK400	XbaI SD T7g10 SfiI SfiI HindIII — lac p/o pelB tet (His)6			●		●	●		
pAK500	XbaI SfiI SfiI EcoRI HindIII — pelB tet dHLX (His)5					●	●		●
pAK600	XbaI SfiI SfiI EcoRI HindIII — pelB tet alkaline phosphatase							●	●
pJB12	XbaI SfiI SfiI EcoRI HindIII — pelB tet KDIR * gIII 250-406	●	●		●				
pJB23	XbaI SfiI SfiI EcoRI HindIII — pelB tet (His)6		●			●	●		
pJB33	XbaI SD T7g10 SfiI SfiI EcoRI HindIII — lac p/o pelB tet (His)6		●	●		●	●		

Fig. 3.4 *Overview of pAK/pJB vector series*. The pAK/pJB vector series (see also Fig. 3.5) can be used either for phage display (pAK100 and pJB12) by the strategy outlined in Fig. 3.1, or for the expression of the antibody in a variety of formats. All vectors contain a chloramphenicol resistance cassette (cam^r) and additionally a tetracycline resistance "stuffer" cassette (*tetA* and *tetR*; 2,101 bp), which will be replaced by the antibody gene (the *tet* cassette allows the monitoring of complete *Sfi*I digested vector by plating of transformed DH5α cells on tetracycline plates). Furthermore, these vectors contain the *lacI* repressor gene, a strong upstream terminator (t_{HP}) to avoid read-through and premature expression, the *lac* promoter/operator and the *pelB* (pectate lyase gene of *Erwinia carotovora*) leader sequence (modified to contain a *Sfi*I site) as well as a

to one RNeasy spin column and continue according to the manufacturer's instructions.

Note: RNA in harvested tissue is not protected from degradation until the sample is mixed with TRIzol Reagent, flash-frozen or disrupted and homogenized in the presence of RNase-inhibiting or protein-denaturing reagents. Therefore, proceed with this step as fast as possible. Generally, DNase digestion is not required with RNeasy Kits since its silica membrane efficiently removes most of the DNA. However, if desired, residual DNA can be removed by optional on-column DNase digestion using the RNase-Free DNase Set (Qiagen). It is important not to overload the RNeasy spin column, as this will significantly reduce RNA yield and quality.

3. Elute the purified RNA by the addition of 30 µl RNase-free water. The mRNA solution is now ready for cDNA synthesis or can alternatively be stored at −80°C for up to one month.

   Note: Diethylpyrocarbonate (DEPC)-treated UHP water can also be used. However, as DEPC is a suspected carcinogen the use of filtrated RNase-free water is recommended.

4. For reverse transcription, take approximately 0.1–0.5 µg RNA and 1 µl random hexamer primers provided in the kit in 20 µl total reaction volume. The precise procedure is described in the QuantiTect Reverse Transcription Kit (Qiagen).

   Note: Ribonuclease H activity of Quantiscript Reverse Transcriptase specifically degrades only the RNA in RNA:DNA hybrids, but it has no effect on pure RNA. Hence, an additional RNA degradation step using another RNase H enzyme is not necessary prior to subsequent PCR reaction.

   Note: Specific primers hybridizing to the constant regions can be used as well, e.g., if only a particular antibody class should be amplified from spleen cells. In general, however, the random hexamer primers work robustly.

Fig. 3.4 (continued) downstream terminator (t_{lpp}). The rationale for these elements has been described in detail previously (Krebber et al. 1997). The origins for phage replication and plasmid replication are as described in Ge et al. (1995). The antibody gene is fused in frame either to $gIII_{250–406}$ for phage display, to a his tag for IMAC purification (Lindner et al. 1992) and *C*-terminal detection with a recombinant anti-his tag scFv-phosphatase fusion protein (Lindner et al. 1997), to dimerization helices (Pack et al. 1993, Plückthun and Pack 1997, see also chapter 7) or to alkaline phosphatase for both dimerization and direct detection (Lindner et al. 1997). In pAK100, the in-frame fusion contains a myc tag (Munro and Pelham 1986), offering an additional detection possibility next to the short *N*-terminal 4-amino acid FLAG tag (DYKD; Knappik and Plückthun 1994) present in all the vectors being encoded by the primers shown in Fig. 3.2. The plasmid pJB12 contains a trypsin cleavage site (KDIR) and can therefore be conveniently used for selection of high-affinity binders as described by Dziegiel et al. (1995) and Johansen et al. (1995). The asterisk in these two vectors pAK100 and pJB12 represents an amber codon. The scFv expression level in pAK400 and pJB33 is enhanced due to the strong Shine Dalgarno sequence SDT7g10 (from T7 phage gene 10). Because of the compatibility of the vectors, this feature can easily be introduced in all of them. In the pJB vector series the co-expressed periplasmic protein Skp (Bothmann and Plückthun 1998), encoded on this vector, can increase the functional yield of antibody fragments expressed in the periplasm without the need of cotransformation with another plasmid coding for further chaperones. This feature can also be introduced into any of the other vectors. The complete sequences of all vectors are available from the authors upon request

3.3.2 PCR Amplification and scFv Assembly

3.3.2.1 PCR Amplification of V_L and V_H Domains

1. Use the primers described in Fig. 3.2, which have been dissolved in 100 µM stock solutions in either sterile water or sterile TE buffer to prepare appropriate mixtures (VL-for mix, VL-rev mix, VH-for mix, and VH-rev mix). Mix them according to the degree of degeneration, indicated as "d" in Fig. 3.2 (equaling the number of different unique sequences encoded by mixed bases in the primer) by adding the stated volumes (in µl) towards the final primer mix. The fraction of lambda-specific primers in both the forward and reverse V_L mixture amounts for ~5% of the total volume, accounting for the low percentage of this light chain type in mouse antibodies. The nominal total primer concentration of these mixtures is still 100 µM, ranging from 3 to 40 µM for each of the individual oligonucleotides.

   Note: As described in the introduction, problems in the cloning of monoclonal antibodies can occur if the hybridoma transcribes more than one functional or even nonfunctional heavy or light chain variable region gene. Therefore, it is highly recommended to omit any lambda chain-specific primer in the PCR if the isotyping already indicates that the hybridoma of interest secretes IgGs possessing kappa light chains.
2. For PCR amplification of V_L and V_H, use the product of the completed first-strand cDNA reaction and prepare the following mixtures:

PCR mix for amplification of V_L	PCR mix for amplification of V_H
2 µl cDNA	2 µl cDNA
1 µl dNTPs (10 mM each)	1 µl dNTPs (10 mM each)
5 µl 10× ThermoPol buffer (NEB)	5 µl 10× ThermoPol buffer (NEB)
0.5 µl VL-for primer mix (100 µM)	0.5 µl VH-for primer mix (100 µM)
0.5 µl VL-rev primer mix (100 µM)	0.5 µl VH-rev primer mix (100 µM)
2.5 µl DMSO	2.5 µl DMSO
0.5 µl VentR Polymerase 2 U/µl (NEB)	0.5 µl VentR Polymerase 2 U/µl (NEB)
38 µl H_2O	38 µl H_2O

   Note: This standard protocol is optimized for VentR polymerase, a DNA polymerase with a 5–15 fold lower error rate than Taq DNA Polymerase (due to an intrinsic 3′→5′ proofreading exonuclease activity). If using other proofreading polymerases (e.g., Phusion High-Fidelity DNA Polymerase from Finnzymes), reaction and PCR program conditions might have to be adapted. If the proposed PCR mix does not lead to any product, varying the cDNA template amount might be beneficial. If the thermocycler does not have a heated cover, add one drop of mineral oil to the reaction tube to prevent evaporation.

3. Perform the following PCR cycles after an initial denaturation of the DNA template for 3 min at 95°C: 5 cycles of 30 s at 95°C, 30 s initial annealing at 55°C, and 45 s elongation at 72°C, followed by 20 cycles of 30 s at 95°C, 30 s at 63°C, and

45 s at 72°C. After the last cycle is completed, an additional 5 min elongation step at 72°C should be performed before cooling the thermocycler to 4°C.

Note: We recommend using a hot start, keeping the PCR tubes on ice and not placing them into the thermocycler until the block has reached 95°C, to minimize unspecific amplification. For successful amplification of V_L and V_H, complete annealing of the 3′-ends of the primers with the template DNA is essential. The recommended annealing temperature of 55°C should be suitable for approx. 97% of the sequences found in a reference set of murine germline sequences in the IMGT database. However, as it is not clear a priori which somatic mutations a given monoclonal antibody may carry in the primer regions, we recommend using a gradient PCR program (covering a range between 70° and 50°C in steps of 2°) to determine the optimum annealing temperature and to amplify the antibody genes without unspecific secondary bands. Alternatively, the PCR might also be run in a "touchdown" manner (Don et al. 1991), starting at an annealing temperature of 70°C and ending at 50°C. As after 5 cycles the amplified PCR product will serve itself as template DNA, the annealing temperature of the last 20 cycles can be increased to 63°C.

4. Analyze 1/10 volume of each PCR mixtures by agarose gel electrophoresis, purify the V_L and V_H genes using the PCR clean-up Gel Extraction Kit (Macherey Nagel) according to the manufacturer's instructions and determine the DNA concentration of both genes.

   Note: Using the listed primer mixtures, the expected lengths of the PCR products of V_L and V_H are between 375–402 bp and 386–440 bp, respectively. Purification of the PCR products is important to remove any residual primers which might interfere with the subsequent assembly PCR. For the case of multiple bands on the agarose gel, gel-purify the band of correct size using the GenElute Gel Extraction Kit (Sigma-Aldrich). If the final DNA concentration is too low afterward, perform a second PCR using these purified fragments as template for gaining sufficient yields of high-quality DNA.

3.3.2.2 Assembly of V_L and V_H by SOE-PCR (Splicing by Overlap Extension)

1. For the assembly PCR, use approximately 10 ng of the PCR product of both domains in a total volume of 50 µl, containing 200 µM dNTPs, 3–5% DMSO, 1 µM outer-for, and outer-rev primer (each) and 1 unit VentR DNA Polymerase (NEB). Following a 3 min 95°C step, perform 5 cycles of 1 min at 95°C, 1 min at 63°C, and 1 min at 72°C, followed by another 5 cycles of 1 min at 95°C, 30 s at 56°C, and 1 min at 72°C and finally 25 cycles of 1 min at 95°C, 90 s at 72°C.

   Note: Hot start PCR and initial assembly of V_L and V_H in the absence of the primers is usually not necessary but can be performed. It is important to include DMSO in the PCR mix as well as to keep the primer concentration as low as indicated to prevent any risk of primer-dimer formation.

   Note: The assembly, as used here, places V_L in front of V_H. This has the advantage of placing a shortened FLAG tag, consisting of only four amino acids,

at the *N*-terminus of the construct. Since its last amino acid, Asp, is the same as the first residue of the V_L domain, only three additional amino acids are needed (Knappik and Plückthun 1994) for allowing specific detection using this tag. A slight asymmetry in the V_H/V_L heterodimer with respect to the pseudo two-fold axis (Plückthun et al. 1996) is taken care of with a 20-amino acid linker, leading to monomeric scFv fragments.

3.3.3 Digestion and Cloning of scFv Genes

1. Purify the product of the assembly PCR using the PCR clean-up Gel Extraction Kit (Macherey Nagel) according to the manufacturer's instructions, eluting the product in 30 µl of the recommended buffer. In case there are several bands on the analytical agarose gel, carry out a gel purification of the correct band, as described in 3.3.3.4.
2. Perform a *Sfi*I digest of the amplified scFv gene for 3–4 h at 50°C (At 37°C, the activity of *Sfi*I would be 10 fold-lower). To the 30 µl purified PCR product, add 5 µl 10× NEbuffer 4 (NEB), 5 µl 10× BSA (final concentration, 100 µg/µl), 9 µl H_2O, and 1 µl (=20 units) *Sfi*I (NEB).
3. Digest appropriate amounts of vector (pAK100 or pJB12, see Fig. 3.4) with *Sfi*I in the presence of the above-mentioned buffer, including BSA. Use 10 units *Sfi*I for 1 µg vector in 50 µl and incubate 4 h at 50°C. Dephosphorylate the cut vector by adding Calf Intestinal Alkaline Phosphatase (CIP, NEB; 0.5 unit/µg vector) to the digestion mix after 2 h and continue incubation for another 2 h at 50°C.

   Note: Dephosphorylation should not be necessary because of the asymmetric overhangs. However, we always include this step to eliminate any risk of religation of single-cut vector.

   Note: pAK100 or pJB12 are phage display vectors (Fig. 3.4). When starting from hybridomas, one can also directly clone the V_L and V_H genes into an scFv expression vector with a stronger promoter, such as pAK400, which does not encode a fusion with gIII. However, depending on the number of additional V genes expressed in the hybridoma, a large number of clones may have to be screened from individual colonies.
4. Purify the digested scFv antibody genes and vector by preparative agarose gel electrophoresis in combination with the GenElute Gel Extraction Kit (Sigma-Aldrich).

   Note: For obtaining pure preparations of a fully digested vector, it is very important not to overload the agarose gel. Furthermore, the gel electrophoresis has to be run long enough to separate small amounts of undigested vector from the digested vector band. For large-scale vector or insert preparation, electro-elution might be an efficient and convenient alternative. If the concentration of eluted DNA is too low for further applications, Microcon 30 columns (Amicon) can be used for concentration.

5. Ligate 50 ng scFv gene fragment with the vector (molar ratio of vector to insert 1:5) with 5 units T4 DNA ligase (NEB) in the presence of 1× T4 DNA ligase buffer in 10 µl volume. Incubate for 2 h at room temperature or overnight at 16°C.

   Note: The ATP-concentration is very crucial for the successful ligation by T4 DNA ligase. Therefore, we recommend using T4 DNA ligase buffer aliquots, which have been properly stored at −20°C and not thawed repeatedly. To allow an easy subcloning of the scFv fragment into vectors for optimized soluble expression and other purposes, compatible vector sets are available (Figs. 3.4 and 3.5).
6. Transform 50 µl chemocompetent XL1-Blue cells (Stratagene) with 5 µl of the ligation mix by heat-shock for 45 s at 42°C, add 500 µl of 2× YT medium after 2 min incubation on ice, and incubate for 45 min, shaking at 37°C.

   Note: Make sure not to exceed a ratio of ligation mix/cells of 1:10 (v/v). Chemocompetent *E. coli* are used, if only a very small diversity of clones is expected, e.g., when cloning from a hybridoma. If a larger diversity and thus many clones are required (e.g., when cloning from spleen cells), follow the instructions for electroporation described in steps 3.3.5.1–3.3.5.3.
7. Plate the transformed cells on 2× YT, 1% glucose, chloramphenicol (30 µg/ml) agar plates, and incubate overnight at 37°C.

   Note: You may check the ratio of desired ligation product to background by including transformation with "religated" plasmid in the absence of any insert. Alternatively, the background signal can be analyzed by testing for tetracycline resistance after transformation of other *E. coli* strains not possessing an intrinsic *tet* resistance (like Invitrogen's DH5α) with the ligation mix. The portion of vector with unremoved or religated *tet* cassette is typically in the range of 0.01–0.1%.

3.3.4 Preparation of Electrocompetent E. coli

1. For preparation of electrocompetent *E. coli* XL1-Blue cells (Stratagene), use 2 ml of a dense overnight pre-culture to inoculate 500 ml medium (2× YT, 15 µg/ml tetracycline). Shake it at 25°C until an OD_{600} of 0.6 is reached, then chill the culture on ice as quickly as possible for 30 min (cool the whole shake flask in a large ice bath).

   Note: Sufficient agitation during growth seems to be very important for preparation of electrocompetent cells, reaching reproducible efficiencies of 3–6 × 10^9 cfu/µg pUC19 DNA. Therefore, use 5 l baffled shake flasks with only 500 ml medium and make sure that the amplitude of the shaker is high enough to vigorously circulate the medium.

   Note: The use of electrocompetent bacteria is an alternative to 3.3.3.6 and needed when a large diversity is expected, typically when cloning from spleen cells.

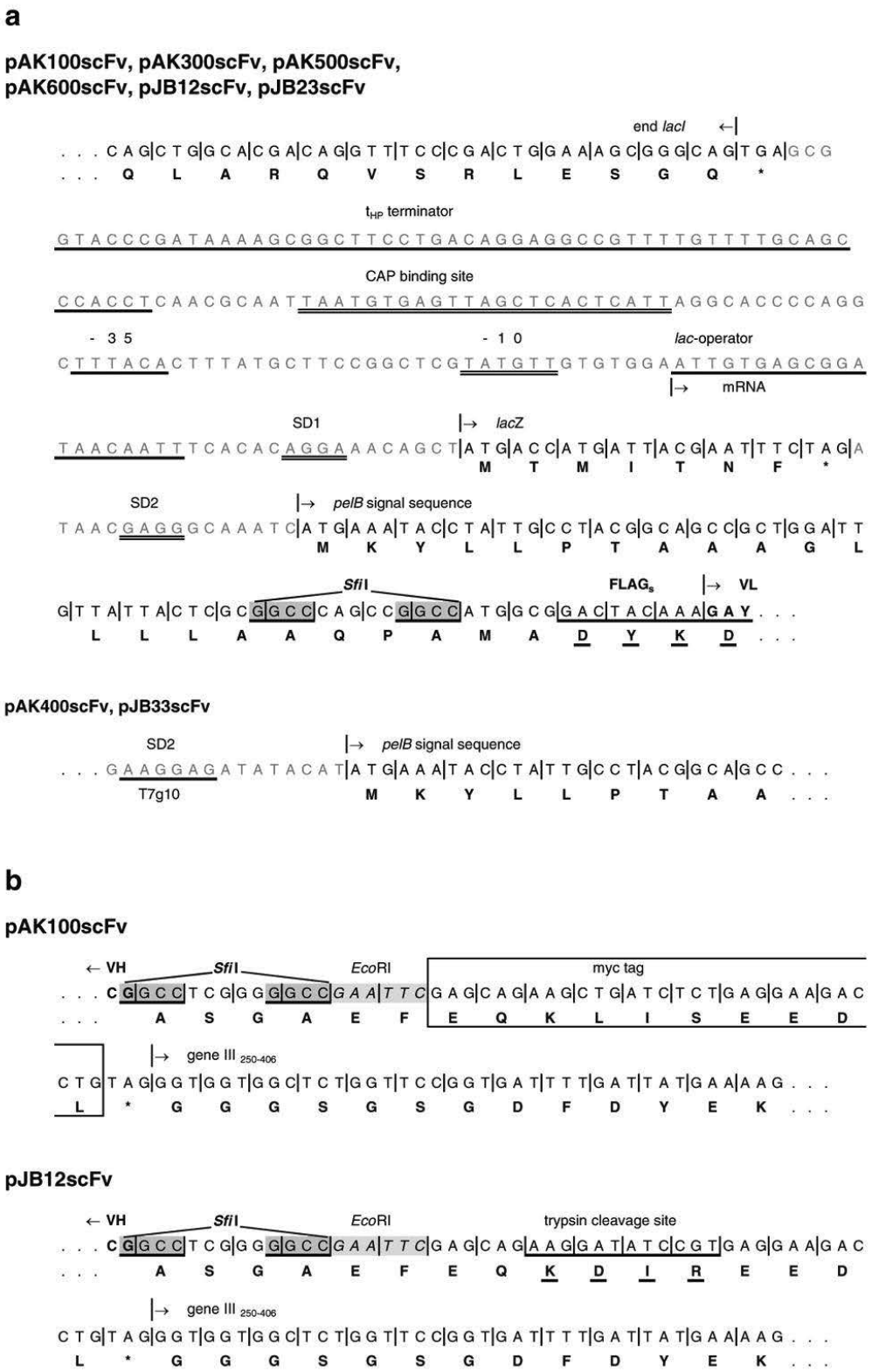


Fig. 3.5 (Continued)

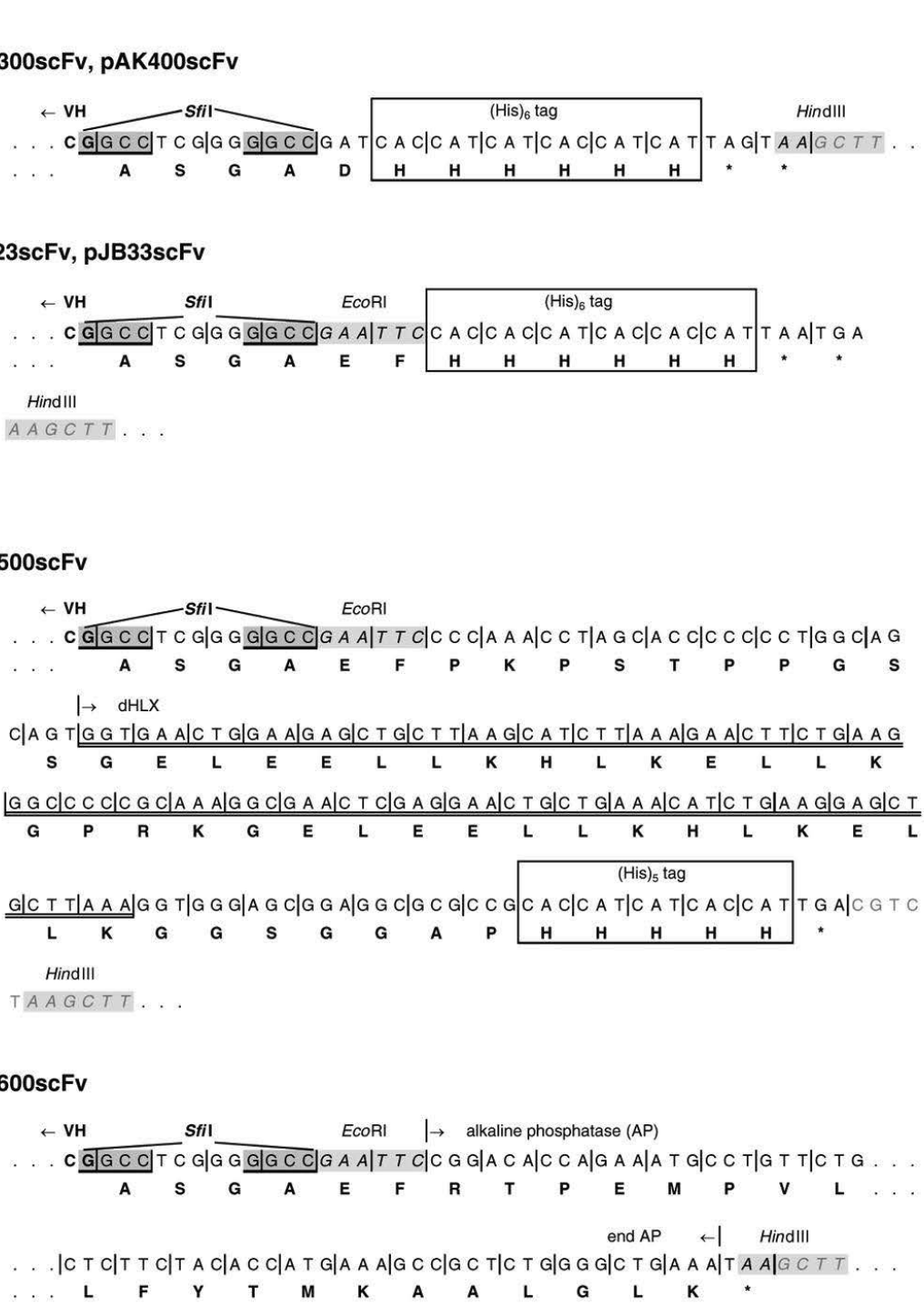


Fig. 3.5 *Detailed sequences upstream and downstream of scFv clonig site*. (**a**) Upstream sequence of pAK100scFv, pAK300scFv, pAK400scFv, pAK500scFv, pAK600scFv, pJB12scFv, and pJB23scFv. The symbol scFv indicates that the vectors are shown after an scFv has been introduced, replacing the *tet* stuffer fragment. The region from the end of the *lacI* repressor gene to the beginning of the antibody V_L domain is shown. The *lacI* repressor gene, t_{HP} terminator sequence, CAP binding site, *lac* operator region, including the -35 and -10 sequence, Shine-Dalgarno (SD) sequence of *lacZ* (SD1), *lacZ* peptide, a second SD sequence (SD2), *pelB* signal

2. Centrifuge the bacterial culture in 50 ml aliquots in disposable tubes for 5 min at 5,000 *g*. Remove as much supernatant as possible (leave the tube upside down for 15–30 s on a clean tissue). Then, fill each tube with 1 volume of ice-cold distilled water (i.e., the same volume as the original culture aliquot) and remove the water immediately (the cell pellet is very solid after this first centrifugation step and will not be resuspended by the brief rinsing with distilled water).

   Note: All these steps should be carried out using ice-cold solutions and be performed in the cold room. Use only ultra pure water to wash cells and to prepare 10% glycerol, as the presence of impurities such as salts in the water might cause the subsequent transformations to fail.
3. Fill each tube with 1 vol distilled water (i.e., the same volume as the original culture aliquot), resuspend the pellet carefully and incubate for 10 min on ice.

   Note: Make sure that the cells are sufficiently solubilized to yield a homogeneous suspension. Cells are best resuspended by swirling rather than pipetting. Never vortex the cell suspension!
4. Transfer the cells into new 50 ml tubes and centrifuge at 5,000 *g* for 10 min. Carefully remove the supernatant and resuspend the pellets each in 50 ml prechilled 10% (v/v) glycerol (Fluka). Incubate on ice for 10 min.
5. Centrifuge resuspended cells at 5,000 *g* for 15 min and remove the supernatant (you might lose a small portion of cells – do not put the tubes upside down on tissue in this step!). Carefully resuspend the cells in 1/500 of the original culture volume (= 1 ml) 10% (v/v) glycerol, freeze the cells in 100 µl aliquots by dipping the tubes immediately into liquid nitrogen and store them at −80°C.

   Note: Electrocompetent cells can be kept at −80°C for up to 12 months.
6. To determine the transformation efficiency, add 1 µl of 10 pg/µl pUC19 DNA (in water) to 40 µl of barely thawed cells (see step 3.3.5.2). Fifty colonies per 1/1,000 of the transformation volume plated correspond to an efficiency of 5×10^9 cfu/µg pUC19 DNA.

Fig. 3.5 (continued) sequence, *N*-terminal *Sfi*I site (*underlined and highlighted*), four amino acid $FLAG_s$ tag (*underlined*), and the start of the V_L domain (sequence GAY; *bold*) are indicated above the sequence. In addition, also the corresponding amino acid sequence is shown. In pAK400 and pJB33, the 15 bp upstream from the *pelB* start codon are replaced by another sequence, including the SD sequence of the phage T7 gene10, while everything else is identical. Because of the modularity of the vectors, this feature can be easily introduced into any of the other vectors (see Fig. 3.4) (**b**) Downstream sequence of pAK100scFv and pJB12scFv. The last two bases of V_H (*bold*), the *Sfi*I and *Eco*RI restriction sites, myc tag (*boxed*) or trypsin cleavage site and the start of $geneIII_{250–406}$ are indicated above the sequence. Asterisks indicate amber stop codon, leading to scFv-gene III fusions upon expression in *E. coli* strains with amber suppressor tRNA, such as XL1-Blue. (**c**) Downstream sequence of pAK300scFv, pAK400scFv, pJB23scFv, and pJB33scFv. The last two bases of V_H (*bold*), the *Sfi*I and *Eco*RI restriction sites and $(His)_6$ tag (*boxed*) are indicated above the sequence. (**d**) Sequences of the downstream *Eco*RI/*Hin*dIII fusion cassettes as used in pAK500 and pAK600. The dHLX dimerization motif (*double underlined*) was taken from Pack et al. (1993). The complete sequence of the mature *E. coli* alkaline phosphatase (AP) gene can be found in Shuttleworth et al. (1986). For the *Eco*RI/*Hin*dIII cloning cassette the two internal *Eco*RI sites of the AP gene have been removed by silent mutations. The complete sequences of all vectors are available from the authors upon request

3.3.5 Library Preparation/Construction

1. For desalting the DNA prior to electroporation, apply the ligation mix to StrataClean Resin (Stratagene; hydroxylated silica, binding proteins with a high affinity, while having a low affinity for DNA at near neutral pH), followed by precipitation in 70% ethanol. Since most salts and small organic molecules are soluble in 70% ethanol, they can be separated from DNA by centrifugation. Resuspend the precipitated DNA in ultra pure water.
2. For each transformation, use desalted ligation mixtures corresponding to 20–100 ng insert. Add the DNA to 40 µl of barely thawed cells on ice and mix by flipping the tube shortly and gently. Immediately transfer the cell-DNA mix to chilled electroporation cuvette (bubble free), pulse according to the guidelines of the electroporator's manufacturer, and add 1 ml of SOC medium (20 g/l bacto-tryptone, 5 g/l yeast extract, 10 mM NaCl, 2.5 mM KCl, 20 mM glucose, 10 mM $MgCl_2$, 10 mM $MgSO_4$) to cells immediately after the pulse.

   Note: For efficient transformation ($\geq 10^8$ clones per µg insert DNA), the time constant using 2 mm cuvettes should be $\geq$5 ms, reflecting properly washed cells. Also, make sure that no air-bubbles are trapped in the cell-DNA mix as they will interfere with the electroporation.
3. Resuspend cells completely in SOC medium and shake for 1 h at 37°C. Afterward, plate dilutions on 2× YT, 1% glucose, chloramphenicol (30 µg/ml) agar plates. Use a sterile spreader or sterile glass beads to evenly distribute the culture over the surface of the 12 × 12 cm plate (do not exceed about 5,000 clones per square plate) and incubate overnight at 37°C. The next day, scrape the colonies off the plates in 3–4 ml 2× YT, containing 30% glycerol, and subsequently store them at −80°C.

   Note: Take care that your library is homogeneously mixed.
4. For phage panning as described in 3.3.6, inoculate cultures with at least tenfold more viable cells than colonies obtained after transformation, in order to have sufficient oversampling. When starting from spleen cells, perform three rounds of phage panning as described in 3.3.6 before testing single clones. When starting from a hybridoma, one round should be sufficient, and in ideal cases, single clones can be tested right away by phage ELISA.

   Note: The first panning round is the most crucial, as you might lose any desired, but less abundant antibody sequence by too extensive washing. Therefore, do not exceed ten washing steps in this first panning round. The panning procedure is analogous to the phage ELISA (3.3.6), except that a pool of phages is grown and that phages are eluted from antigen (at the end of 3.3.6.4), which are afterward added again to exponentially growing bacteria. This is described in detail elsewhere (Barbas et al. 2001; Lee et al. 2007) and also in this volume.

   Note: The screening of single clones can be performed in three ways. First, at the level of phages (phage ELISA), as described in Sect. 3.3.3.6, second, after retransforming of an suppressor tRNA-deficient strain such as, e.g., *E. coli* strain JM83 (Yanisch-Perron et al. 1985), still with the amber codon, containing

pAK100 derived plasmids, or third, after recloning into a more efficient expression vector such as pAK400 (Fig. 3.5), which carries no gene III. In second and third option, the soluble scFv is screened by ELISA.

Note: This protocol does not describe the periplasmic expression of scFv fragments in *E. coli* (Glockshuber et al. 1990) and their subsequent purification (reviewed in Plückthun et al. 1996). More details can be found in chapter 27 "Improving expression of scFv fragments by coexpression of periplasmic chaperones" in this volume.

3.3.6 Screening for Binders by Phage ELISA

1. When starting from hybridoma, pick 10 colonies (from spleen, as many as you can handle) and grow them separately at 37°C in 2 ml 2× YT, 1% glucose, chloramphenicol (30 μg/ml) until they reach an OD_{600} of 0.5. This level of glucose fully represses expression, and, thus, the growth temperature can be 37°C. Dilute 1:10 in 2× YT, 1% glucose, chloramphenicol (30 μg/ml), containing 1 mM IPTG, and 1×10^{10} pfu VCSM13 helper phage (Stratagene) per ml, and grow overnight at 26°C or 37°C (for some murine scFvs with aggregation tendencies, growth at 26°C after infection may be necessary). The phage titer after overnight incubation is in the range of 10^{11}–10^{12} cfu per ml supernatant.

   Note: XL1-Blue should be grown on agar plates and in media containing tetracycline (*tet*) as the F′-plasmid encoding for the F-pili required for infection of bacteria also carries the *tet* resistance gene. The phage titer (in cfu) should be determined in order to rule out any problems during phage production. To do so, take a log-phase culture of XL1-Blue cells (OD_{600} = 0.4–0.6) and incubate aliquots of this culture with serial dilutions of your phage preparation. After 15 min incubation at 37°C, plate appropriate amounts (30–150 cfu/plate) on 2× YT, 1% glucose, chloramphenicol (30 μg/ml) agar plates.

   Note: We are aware of the fact that the presence of such a high level of glucose during IPTG-induction is rather unusual. However, for phage display, the induction level does not have to be very high. Based on our experiences with this vector series described here, a combination of IPTG addition and the presence of glucose seems to be crucial for the successful expression of some scFv fragments, notably those with nonideal biophysical properties, and appropriate for most, but may have to be checked for each scFv individually in case of unusual properties.
2. Centrifuge the culture 10 min at 16,000 *g* and 4°C. Take 1.6 ml supernatant and mix it with 0.4 ml 20% PEG 6000 (Fluka), 2.5 M NaCl in a 2 ml Eppendorf tube in order to precipitate the phages (Sambrook and Russell 2001).

   Note: We recommend that the PEG solution be freshly prepared.
3. Incubate on ice for 30–60 min and centrifuge for 15 min at 5,600 *g* and 4°C.

   Note: It is important not to centrifuge phages at too high a *g* force, as otherwise, it will be difficult to resuspend them homogeneously, resulting in a

decreased phage titer. The size of the white pellet does not necessarily reflect a high or low phage titer.

4. Resuspend the phage pellet in 400 µl PBS (with 10% (v/v) glycerol). For complete resuspension, incubate the phage solution on an orbital shaker at 800 rpm for 15 min at 4°C. Pellet insoluble matter (cell debris) by centrifugation for 10 min at 11,000 *g* and 4°C and transfer the phage solution to a fresh tube. Use 100 µl phage solution per well in an ELISA assay to distinguish phages displaying functional scFv antibody from those which display nonfunctional or nonproductive antibody fragments.
5. If soluble antigen is available, include a competition ELISA control showing that free antigen is able to compete with bound antigen for phage binding to distinguish nonspecific "sticky" from specifically binding phages. In principle, the same ELISA protocol that was used for the hybridoma screening procedure can be used.

   Note: For weak binders, it might be important to use more phages for ELISA analysis. In this case, the culture volume should be increased ten times. If no functional clone shows up in ELISA of single clones, perform one round of phage panning in order to enrich the functional binders. The enrichment should be checked by comparison of eluted phages from a specific surface versus a surface without antigen. In addition, it is recommended to analyze the phage solution by immunoblot, using an anti-M13 HRP-conjugated antibody (GE Healthcare), to ensure the correct fusion of the scFv to the gIII-protein as well as its correct display on the phage surface.

3.4 Troubleshooting

This part of the protocol contains general comment about potential pitfalls of the recommended standard method. The most critical steps were already highlighted directly following the instructions in the different subsections.

(a) In case of low transformation yields, check whether the problem is the transformation itself or rather the ligation. To investigate the quality of ligation, analyzing an aliquot by agarose gel electrophoresis might indicate any problems caused by nucleases. Furthermore, it might be informative to compare the ligation efficiency of *Sfi*I digested PCR product with inserts derived from plasmid digestion. In order to check both the ligation and the transformation efficiency, a defined amount of pUC19 DNA can be added to the ligation mixture. Because of the chloramphenicol resistance of the cloning vector and the ampicillin resistance of pUC19 DNA, it is possible to calculate the ligation efficiency by plating double transformed cells on ampicillin or chloramphenicol plates, respectively, and comparing the number of clones. The transformation efficiency (in presence of the ligation mixture) can be judged by comparison of the colony number after transformation with pUC19 DNA alone.

(b) The quality of the oligonucleotides used in this procedure is crucial for the successful and reliable amplification of various antibody genes as well as their subsequent assembly into scFvs. The number of proposed primers is important for a broad representation of the immune response, as any sequence absent from the complex mixture will obviously decrease the functional library size. We also strongly recommend using primers that have been accurately purified after their synthesis (either by HPLC or, for longer primers, by PAGE) to ensure that no single-base deletions are present in any of the oligonucleotides. These deletions as well as any insertions would cause frameshifts in the final gene assembly, resulting in a number of nonfunctional library members. Therefore, we also suggest – especially for library cloning – to sequence the genes of several random clones as well as to check for full-length scFv by western blot analysis detecting its fusion partner gene III (see note at 3.3.6.5).

(c) In case of severe problems in the PCR amplification of the V_H and V_L genes (steps 3.3.2), it might be worth considering to divide this reaction into two separate ones. Using the proposed primers without any overhang at their 5′-end (which either codes for the $FLAG_s$-tag and the *Sfi*I cleavage site, or the $(Gly_4Ser)_4$-linker), the pure antibody DNA should be amplified in a first PCR reaction, and, subsequently, a second PCR should be performed for reamplification and introduction of the appropriate overhangs with the original full-length primers. This procedure increases the degree of matching in both reactions, and might therefore help in the annealing step of the primers.

(d) Whenever expression of the scFv gene is not required, the bacteria should be grown in the presence of 1% glucose. Glucose will cause a tight suppression of the *lac* promoter, thereby ensuring the genetic stability of the inserted scFv genes. Likewise, we suggest growing XL1-Blue always on agar plates and in media containing tetracycline (*tet*) to keep the bacteria infective, as the *tet* resistance is located on the F′-plasmid that also contains the genes encoding F-pilus formation. Always use fresh XL1-Blue colonies, as subcloning might occasionally lead to the formation of *tet* resistant cells, which are no longer infectable. As the F-pilus expression is reduced when the bacteria are past log phase as well as when grown at temperatures below 34°C, we also recommend growing them at 37°C to OD_{600} = 0.4–0.6.

(e) When working with libraries, double transformants can and will occur (Goldsmith et al. 2007). It is thus highly recommended that the scFv fragments of interest be recloned into a new vector (e.g., from pAK100 to pAK400), when they are analyzed at the level of pure unfused protein, thereby also introducing a stronger translation initiation region. It should be noted that diluted retransformation cannot resolve plasmid mixtures, as in *E. coli,* plasmids can form reversible concatamers.

Acknowledgements This protocol has evolved over the years, and heavily relies on the original versions developed by Anke Krebber and Jörg Burmester, with important discussions and contributions to the reagents and procedures gradually added from Peter Lindner, Lutz Jermutus, Jörg Willuda, Daniel Steiner, Barbara Klinger, and Cornelia Rinderknecht.

References

Adams GP, McCartney JE, Tai MS, Oppermann H, Huston JS, Stafford WF 3rd, Bookman MA, Fand I, Houston LL, Weiner LM (1993) Highly specific in vivo tumor targeting by monovalent and divalent forms of 741F8 anti-c-erbB-2 single-chain Fv. Cancer Res 53:4026–4034

Amstutz P, Binz HK, Zahnd C, Plückthun A (2006) Ribosome display: in vitro selection of protein-protein interactions. In: Celis J (ed) Cell biology: a laboratory handbook, 3rd edn. Elsevier, Amsterdam, pp 497–509

Backmann N, Zahnd C, Huber F, Bietsch A, Plückthun A, Lang HP, Güntherodt HJ, Hegner M, Gerber C (2005) A label-free immunosensor array using single-chain antibody fragments. Proc Natl Acad Sci USA 102:14587–14592

Barbas C, Burton D, Scott J, Silverman G (2001) Phage display: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Bothmann H, Plückthun A (1998) Selection for a periplasmic factor improving phage display and functional periplasmic expression. Nat Biotechnol 16:376–380

Burmester J, Plückthun A (2001) Construction of scFv fragments from hybridoma or spleen cells by PCR assembly. In: Kontermann R, Dübel S (eds) Antibody engineering. Springer, Berlin, pp 19–40

Di Paolo C, Willuda J, Kubetzko S, Lauffer I, Tschudi D, Waibel R, Plückthun A, Stahel RA, Zangemeister-Wittke U (2003) A recombinant immunotoxin derived from a humanized epithelial cell adhesion molecule-specific single-chain antibody fragment has potent and selective antitumor activity. Clin Cancer Res 9:2837–2848

Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res 19:4008

Dziegiel M, Nielsen LK, Andersen PS, Blancher A, Dickmeiss E, Engberg J (1995) Phage display used for gene cloning of human recombinant antibody against the erythrocyte surface antigen, rhesus D. J Immunol Methods 182:7–19

Ewert S, Honegger A, Plückthun A (2004) Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering. Methods 34:184–199

Freund C, Ross A, Plückthun A, Holak TA (1994) Structural and dynamic properties of the Fv fragment and the single-chain Fv fragment of an antibody in solution investigated by heteronuclear three-dimensional NMR spectroscopy. Biochemistry 33:3296–3303

Ge L, Knappik A, Pack P, Freund C, Plückthun A (1995) Expressing antibodies in Escherichia coli. In: Borrebaeck C (ed) Antibody engineering, 2nd edn. Oxford University Press, Oxford, pp 229–236

Glockshuber R, Malia M, Pfitzinger I, Plückthun A (1990) A comparison of strategies to stabilize immunoglobulin Fv-fragments. Biochemistry 29:1362–1367

Goldsmith M, Kiss C, Bradbury AR, Tawfik DS (2007) Avoiding and controlling double transformation artifacts. Protein Eng Des Sel 20:315–318

Hanes J, Plückthun A (1997) *In vitro* selection and evolution of functional proteins by using ribosome display. Proc Natl Acad Sci USA 94:4937–4942

Hanes J, Jermutus L, Weber-Bornhauser S, Bosshard HR, Plückthun A (1998) Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries. Proc Natl Acad Sci USA 95:14130–14135

Honegger A, Plückthun A (2001) The influence of the buried glutamine or glutamate residue in position 6 on the structure of immunoglobulin variable domains. J Mol Biol 309:687–699

Honegger A, Spinelli S, Cambillau C, Plückthun A (2005) A mutation designed to alter crystal packing permits structural analysis of a tight-binding fluorescein-scFv complex. Protein Sci 14:2537–2549

Huston JS, Levinson D, Mudgett-Hunter M, Tai MS, Novotny J, Margolies MN, Ridge RJ, Bruccoleri RE, Haber E, Crea R, Oppermann H (1988) Protein engineering of antibody binding

sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. Proc Natl Acad Sci USA 85:5879–5883

Johansen LK, Albrechtsen B, Andersen HW, Engberg J (1995) pFab60: a new, efficient vector for expression of antibody Fab fragments displayed on phage. Protein Eng 8:1063–1067

Jung S, Spinelli S, Schimmele B, Honegger A, Pugliese L, Cambillau C, Plückthun A (2001) The importance of framework residue H6, H7, and H10 in antibody heavy chains: Experimental evidence for a new structural subclassification of antibody V_H domains. J Mol Biol 309:701–716

Kabat E, Wu T, Reid-Miller M, Perry H, Gottesman K, Foeller C (1991) Sequences of proteins of immunological interest, 5th edn. US Department of Health and Human Services, Public Service, NIH

Knappik A, Plückthun A (1994) An improved affinity tag, based on the FLAG peptide for the detection and purification of recombinant antibody fragments. Biotechniques 17:754–761

Knappik A, Ge L, Honegger A, Pack P, Fischer M, Wellnhofer G, Hoess A, Wölle J, Plückthun A, Virnekäs B (2000) Fully synthetic human combinatorial antibody libraries (HuCAL), based on modular consensus frameworks and CDRs randomized with trinucleotides. J Mol Biol 296:57–86

Krebber A, Bornhauser S, Burmester J, Honegger A, Willuda J, Bosshard HR, Plückthun A (1997) Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. J Immunol Methods 201:35–55

Kütemeier G, Harloff C, Mocikat R (1992) Rapid isolation of immunoglobulin variable genes from cell lysates of rat hybridomas by polymerase chain reaction. Hybridoma 11:23–32

Lee CM, Iorno N, Sierro F, Christ D (2007) Selection of human antibody fragments by phage display. Nat Protoc 2:3001–3008

Lefranc M, Lefranc G (2001) The Immunoglobulin FactsBook. Academic Press, London

Lindner P, Guth B, Wülfing C, Krebber C, Steipe B, Müller F, Plückthun A (1992) Purification of native proteins from the cytoplasm and periplasm of Escherichia coli using IMAC and histidine tails: a comparison of proteins and protocols. Methods 4:41–56

Lindner P, Bauer K, Krebber A, Nieba L, Kremmer E, Krebber C, Honegger A, Klinger B, Mocikat R, Plückthun A (1997) Specific detection of his-tagged proteins with recombinant anti-His tag scFv-phosphatase or scFv-phage fusions. Biotechniques 22:140–149

Luginbühl B, Kanyo Z, Jones RM, Fletterick RJ, Prusiner SB, Cohen FE, Williamson RA, Burton DR, Plückthun A (2006) Directed evolution of an anti-prion protein scFv fragment to an affinity of 1 pM and its structural interpretation. J Mol Biol 363:75–97

Mondon P, Dubreuil O, Bouayadi K, Kharrat H (2008) Human antibody libraries: a race to engineer and explore a larger diversity. Front Biosci 13:1117–1129

Morfill J, Blank K, Zahnd C, Luginbühl B, Kuhner F, Gottschalk KE, Plückthun A, Gaub HE (2007) Affinity-matured recombinant antibody fragments analyzed by single-molecule force spectroscopy. Biophys J 93:3583–3590

Munro S, Pelham HR (1986) An Hsp70-like protein in the ER: identity with the 78 kD glucose-regulated protein and immunoglobulin heavy chain binding protein. Cell 46:291–300

Pack P, Kujau M, Schroeckh V, Knüpfer U, Wenderoth R, Riesenberg D, Plückthun A (1993) Improved bivalent miniantibodies, with identical avidity as whole antibodies, produced by high cell density fermentation of Escherichia coli. Biotechnology (N Y) 11:1271–1277

Plückthun A, Pack P (1997) New protein engineering approaches to multivalent and bispecific antibody fragments. Immunotechnology 3:83–105

Plückthun A, Krebber A, Krebber C, Horn U, Knüpfer U, Wenderoth R, Nieba L, Proba K, Riesenberg D (1996) Producing antibodies in Eschericia coli: Fom PCR to fermentation. In: McCafferty J, Hoogenboom H (eds) Antibody engineering: a practical approach. IRL press, Oxford, pp 203–252

Sambrook J, Russell D (2001) Molecular cloning. A laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Schaffitzel C, Zahnd C, Amstutz P, Luginbühl B, Plückthun A (2005) In vitro selection and evolution of protein-ligand interactions by ribosome display. In: Golemis E, Adams P (eds) Protein-protein interactions: a molecular cloning manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 517–548

Shuttleworth H, Taylor J, Minton N (1986) Sequence of the gene for alkaline phosphatase from Escherichia coli JM83. Nucleic Acids Res 14:8689

Tugarinov V, Zvi A, Levy R, Hayek Y, Matsushita S, Anglister J (2000) NMR structure of an anti-gp120 antibody complex with a V3 peptide reveals a surface important for co-receptor binding. Structure 8:385–395

Vaughan TJ, Williams AJ, Pritchard K, Osbourn JK, Pope AR, Earnshaw JC, McCafferty J, Hodits RA, Wilton J, Johnson KS (1996) Human antibodies with sub-nanomolar affinities isolated from a large nonimmunized phage display library. Nat Biotechnol 14:309–314

Winter G, Griffiths AD, Hawkins RE, Hoogenboom HR (1994) Making antibodies by phage display technology. Annu Rev Immunol 12:433–455

Wörn A, Plückthun A (2001) Stability engineering of antibody single-chain F_v fragments. J Mol Biol 305:989–1010

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119

Zahnd C, Spinelli S, Luginbühl B, Amstutz P, Cambillau C, Plückthun A (2004) Directed in vitro evolution and crystallographic analysis of a peptide binding scFv antibody with low picomolar affinity. J Biol Chem 279:18870–18877

Zahnd C, Amstutz P, Plückthun A (2007) Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target. Nat Methods 4:269–279

3.4 Miniantibodies

Jonas V. Schaefer, Peter Lindner and Andreas Plückthun

in **Antibody Engineering** (Kontermann, R., and Dübel, S., eds) Vol. 2, 2nd edit., 2010,pp. 85-99, Springer Verlag, Berlin Heidelberg, Germany.

Chapter 7
Miniantibodies

Jonas V. Schaefer, Peter Lindner, and Andreas Plückthun

Abbreviations

IPTG	Isopropylthiogalactoside
PBS	Phosphate buffered saline
scFv	Single-chain Fv fragment
tet	Tetracycline
LB	Luria-Bertani media
SB	Super broth media

7.1 Introduction

7.1.1 Motivation

The usual motivation to create a multivalent molecule is to increase its functional affinity (avidity) to a corresponding multimeric antigen structure, which can be a cell surface, a virus surface or a fibrous polymer. Obviously, no increase in affinity to a soluble monomeric antigen can be expected. However, an increased functional affinity to a formally monovalent antigen will usually be observed when it is immobilized on a densely packed surface (e.g. on an ELISA plate or a BIAcore chip). The increased size of the antibody fragment upon multimerization, in conjunction with higher functional affinity, can also lead to improved tumor localization (Hu et al. 1996; Todorovska et al. 2001; Deyev et al. 2003; Kubetzko et al. 2006). Besides causing an avidity gain, bivalent binding might also result in agonistic activity, which may or may not be desired in a particular application.

J.V. Schaefer, P. Lindner, and A. Plückthun (✉)
Biochemisches Institut, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
e-mail: plueckthun@bioc.uzh.ch

It is essential to understand that this avidity gain, which can be quantified (see, e.g. Crothers and Metzger (1972); Plückthun and Pack (1997); Müller et al. (1998b)), is a phenomenon closely linked to geometry. The bivalent molecule must be able to reach two epitopes at the same time. Thus, both the distance and the relative orientation of the binding sites will be crucial. It is, therefore, not the same if the two binding sites are linked head to tail, head to head, or tail to tail, even when connected with flexible linkers.

In the IgG molecule, the two binding sites are far enough apart to often reach epitopes from two adjacent protein molecules. The inherent (approximate) C2 symmetry of the antibody is ideal if the two antigens are also related by an approximate twofold axis. In a membrane, where proteins have some mobility, such an arrangement is very often possible. From such considerations we designed the "miniantibodies" to mimic the geometry of IgG molecules, using components which can be made conveniently in bacteria.

A further advantage of this format is that two different specificities can be combined within one miniantibody, offering numerous applications in biotechnology, diagnostics, and potentially therapy. Such applications can include the cross-linking of two cells or, alternatively, binding to two epitopes on the same cell can increase avidity and possibly selectivity. Finally a bispecific molecule can bring a payload to a cell by one arm binding to the cell, the other to the payload.

Dimeric miniantibodies can also be of interest as capture molecules: when immobilized on plastic support, at least one of the molecule's binding sites usually remains functional, while the other eventually may denature upon binding to the surface. In contrast, monovalent scFv fragments normally lose their antigen binding capability upon immobilization to plastic.

7.1.2 Overview of Multimerization

Three principal concepts exist to multimerize antibody scFv fragments (Plückthun and Pack 1997). The first – being the subject of this chapter – is to connect them, usually via a flexible hinge region, to a module or domain that will itself multimerize. We have termed the resulting constructs "miniantibodies" (Pack and Plückthun 1992), as they recreate the basic flexible disposition of the two binding sites of IgG molecules in a smaller assembly. The ability to multimerize antibody formats other than scFv fragments by using the same strategy is readily apparent.

The second possibility is to shorten the linker between the antibody V_L and V_H domains so that a monovalent fragment cannot form (Holliger et al. 1993; Todorovska et al. 2001), creating so-called dia-, tri- or tetrabodies. The third alternative is to connect two or more scFv fragments linearly with flexible linkers (Kellner et al. 2008). These two latter approaches have also been combined successfully in the past (Kipriyanov et al. 1999). While easy to draw as a cartoon, it must be remembered that domains of natural antibody domains are quite aggregation-prone, especially during folding (i.e. expression) and will in such assemblies also

generate partially folded domains that can lead to inactive and heterogeneous products. It is not secured, therefore, that the desired molecular assembly will actually form in reality in good yield in every case, nor that it will have the prescribed binding properties.

The "miniantibody" concept has been the attempt to create a structure with the same "wing span" as an IgG, but composed of modules which can be readily expressed in *E. coli*. This is achieved by the oligomerization domains which are not directly located downstream of the C-terminus of the scFv fragment but rather separated via a spacer, such as the upper hinge region from murine or human IgG3. This confers rotational freedom and flexibility very similar to that of the Fab-arms of full-length antibodies. As the oligomerization domains in the vectors presented here are encoded in different modular gene cassettes, it is possible to switch formats readily between these multimerized antibody fragments by simple cloning, independent of their structural properties (Willuda et al. 2001). Also, many fusions to other oligomeric modules, to enzymes and other proteins have been previously analyzed regarding their dimerization or multimerization potential (Plückthun and Pack 1997; Zhang et al. 2004).

7.1.3 Miniantibodies

The basic concept for all miniantibody constructs is the fusion of the scFv fragment to an oligomerizing element. In the simplest case, this self-associating peptide domain is an amphipathic, α-helix-forming stretch of amino acids (usually between 16 and 40 residues, see Table 7.2 for details) attached to the scFv via a flexible hinge region, giving both partners enough steric freedom to fold individually. As schematically outlined in Fig. 7.1, this leads to dimeric or tetrameric miniantibodies, depending on the oligomerization motif chosen. Most conveniently, the miniantibodies are expressed in the bacterial periplasm to allow the formation of disulfide bonds in their scFv part. All amphipathic oligomerization helices presented in Table 7.1 were therefore chosen to be compatible with periplasmic folding and the transport through the bacterial membrane, causing no significant problems in the folding and expression of most scFv fragments tested.

7.1.4 Dimeric Miniantibody Constructs

While most methods for the formation of bivalent or bispecific antibody fragments require a significant reconstruction of the format compared to the scFv, the generation of dimeric miniantibodies is simply achieved by adding an oligomerizing sequence to the C-terminus of the scFv fragment. Examples for such self-associating modules are the naturally occurring dimerization helix from the yeast transcription factor GCN4 (Fig. 7.1a) (scFv-ZIP; O'Shea et al. 1991; Dürr et al. 1999), the C_H3/

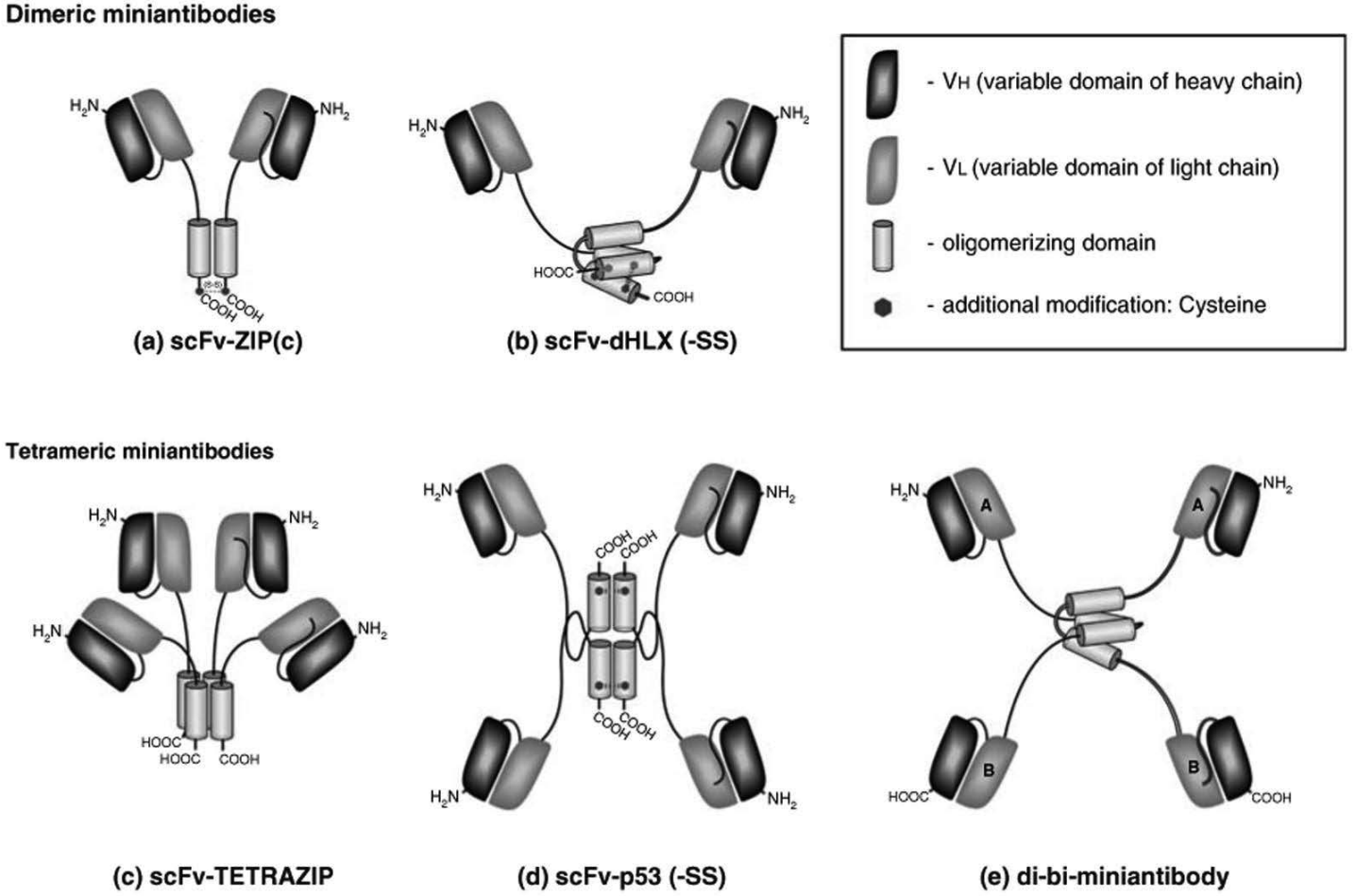


Fig. 7.1 *Schematic representation of oligomeric miniantibody formats*. (**a**) Dimeric GCN4 leucine zipper, scFv-ZIP, the optional disulfide bond is shown in parentheses and with a *dotted line*; (**b**) dimeric helix-turn-helix module, scFv-dHLX, the optional disulfide bonds are shown as *dotted lines*; (**c**) tetrameric modified GCN4 leucine zipper, scFv-TETRAZIP; (**d**) tetramerizing domain of human p53, scFv-p53, the optional disulfide bonds are shown as *dotted lines*; and (**e**) $(scFv)_A$-hinge-dHLX-hinge-$(scFv)_B$ arrangement, di-bi-miniantibody. In each case, V_H and V_L domains of the scFvs are represented in *darker* and

Table 7.1 Cross-references between oligomerizing elements and corresponding plasmids/literature

Construct	Upper hinge	Self-associating peptide	Modifications	Plasmid	Reference
Bivalent					
scFv-ZIP	Murine IgG3	GCN4 leucine zipper	–	pACKZIP	A, B, C, D
scFv-ZIPc	Murine IgG3	GCN4 leucine zipper	C-terminal Cys	pACKZIPc	C, D
scFv-dHLX	Murine IgG3	Helix1-turn-Helix2	–	pAK500	H
scFv-dHLX-SS	Murine IgG3	Helix1-turn-Helix2	Internal Cys	pAK500-SS	J
Bispecific					
scFv-JUN	Murine IgG3	JUN leucine zipper	–	pACKIHJUN	I
scFv-FOS	Murine IgG3	FOS leucine zipper	–	pACKFOS	I
CH1-CL	Murine IgG3	CH1 and CL from IgG	–	pKM30425 M1CHCL	F
Tetravalent					
scFv-TETRAZIP	Murine IgG3	GCN4 leucine zipper, modified	–	pACKtZIP	A, D
scFv-p53	Human IgG3	Oligomerization domain of human p53	–	pMStetp53His	E
scFv-p53-SS	Human IgG3	Oligomerization domain of human p53	Internal Cys	–	J
Tetravalent/bispecific					
di-bi	Murine IgG3	Helix1-turn-Helix2	–	pKM310M1dhlx 425h	G

Important elements of various miniantibody formats are listed as overview. For exact amino acid sequences of the elements, see Table 7.2. Vectors carrying miniantibody genes in these formats and references are given. Letters in the reference column denote: (A) Pack et al. 1995; (B) Pack et al. 1993; (C) Pack and Plückthun 1992; (D) Ge et al. 1995; (E) Rheinnecker et al. 1996; (F) Müller et al. 1998c; (G) Müller et al. 1998b; (H) Krebber et al. 1997; (I) Plückthun and Pack 1997; (J) Kubetzko et al. 2006

Fig. 7.1 (continued) *lighter color*, respectively. Linker and hinge regions are shown as *black lines* (either filled or not, indicating different polypeptide chains within homo- or heterodimeric constructs). The respective helical elements responsible for oligomerization are depicted as grey cylinders. Their orientation is derived from the published crystal structures of the coiled coil (PDB 2zta) (O'Shea et al. 1991), tetrazipper (PDB 1gcl) (Harbury et al. 1993), the NMR structure of the designed dHLX motif (PDB 1qp6) (Hill and Degrado 1998) and both NMR and X-ray structures of the p53 tetramerization domain (PDB 1aie) (Jeffrey et al. 1995; Mittl et al. 1998). Several of the constructs have also been modified with additional cysteines to allow disulfide formation (Kubetzko et al. 2006). This is schematically indicated by *dots* for the cysteines and *dotted lines* for the disulfide bonds. These cysteine modifications are optional, and the expression yield is generally higher when not using the cysteine modified modules

F_c-domains of antibodies (see Chap. 30) or synthetically designed two- or four-helical bundle elements (Eisenberg et al. 1986). The latter motif consists of a helix-turn-helix motif fused to the scFv with two of them "clasping" each other (see scFv-dHLX in Fig. 7.1b).

Bispecific miniantibodies can be created if two different scFvs are chosen and fused to modules forming specific heterodimers. However, as not all heterodimerizing modules also work well in vivo, problems of homodimerization and proteolytic susceptibility have to be taken into consideration. The question of specific heterodimerization was tackled using an in vivo selection approach with different libraries for both helices (Arndt et al. 2000) resulting in coiled coil helices which showed very good behavior with regard to stability, heterospecificity, and resistance to proteases (Arndt et al. 2001).

7.1.5 Tetrameric Miniantibody Constructs

Specific amino acid exchanges in all hydrophobic contact positions a and d of the GCN4 zipper (reviewed by Woolfson 2005) result in the self-assembly of a stable tetrameric bundle (Harbury et al. 1993), and fusing this modified zipper version to a scFv leads to tetrameric miniantibodies (scFv-TETRAZIP in Fig. 7.1c; Pack et al. 1995). A low immunogenicity can be expected for the fusion of a humanized scFv with the tetramerization domain of human p53 (Jeffrey et al. 1995; Rheinnecker et al. 1996), as it also uses a human IgG3 hinge (Table 7.1).

7.1.6 Extensions of the Miniantibody Concept

To further stabilize the multimeric formats, intermolecular disulfide bonds were designed and introduced in variants of the scFv-dHLX and scFv-p53 formats, resulting in the respective SS mutants (see Table 7.2). The presence of the newly introduced cysteine residues (two in the dimerization motif dHLX and one in p53) was shown to cause the formation of covalent cross-links of the self-associated peptides, thus increasing their stability (Kubetzko et al. 2006). On the basis of similar considerations, single cysteine residues can also be added to the C-terminus of the leucine zipper (c variant in Fig. 7.1a), resulting in covalent linkage by disulfide bond formation (Pack and Plückthun 1992). However, it has to be kept in mind that incorrect disulfides may also be formed in these constructs with additional cysteines, leading to slightly lower yields of correctly folded multimeric antibody fragments.

A combination of directed bivalency with bispecificity can be obtained by using so-called "di-bi-miniantibodies" (Müller et al. 1998a). In this construct, a second scFv is fused downstream the dimerization motif, resulting in a $(scFv)_A$-hinge-dHLX-hinge-$(scFv)_B$ arrangement (Fig. 7.1e).

Table 7.2 Amino acid sequences of hinges and oligomerizing elements

Construct	Element	Amino acid sequence
	Murine IgG3 upper hinge	PKPSTPPGSS
	Human IgG3 upper hinge	TPLGDTTHTSG (present in scFv-p53 constructs)
scFv-ZIP	GCN4 leucine zipper	RMKQLEDKVEELLSKNYHLENEVARLKKLVGER
scFv-ZIPc	GCN4 leucine zipper – Cys	RMKQLEDKVEELLSKNYHLENEVARLKKLVGER–GG**C**GG
scFv-dHLX/di-bi	Helix1-turn-helix2 – spacer – $(His)_5$	GELEELLKHLKELLKG-PRK-GELEELLKHLKELLKG–*GSGGAP*–HHHHH
scFv-dHLX-SS	Helix1-turn-helix2 with two internal disulfide bonds	GELEELLKHLKELLKG-PRK-GEL**C**ELLKHLKEL**C**KG–*GSGGAP*–HHHHH
scFv-JUN	JUN leucine zipper	RIARLEEKVKTLKAQNSELASTANMLREQVAQLKQKVMNY
scFv-FOS	FOS leucine zipper	LTDTLQAETDQLEDKKSALQTEIANLLKEKEKLEFILAAH
scFv-TETRAZIP	GCN4 leucine zipper, modified	R**L**KQI**E**DK**L**EE**I**LSK**L**YHI**E**NE**L**AR**I**KK**LL**GER
scFv-p53	Oligomerization domain human p53 – spacer – $(His)_5$	KPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEP–*GGSGGAP*–HHHHH
scFv-p53-SS	Oligomerization domain human p53 with one internal disulfide bond	KPLDGEYFTLQIRGRERFEMFRELNE**C**LELKDAQAGKEP–*GGSGGAP*–HHHHH

The amino acid sequences (one-letter code) of various oligomerizing modules and hinges are given. In the modified GCN4 leucine zipper, which leads to tetramerization, the exchanged amino acids are in *bold-face* as are the cysteine residues introduced for stabilization purposes in other variants. The histidine-tags for detection and purification purposes are *underlined*. The amino acids EF and the end of some constructs were introduced for an *Eco*RI-restriction site. Cross-references to the corresponding vectors and literature are listed in Table 7.1

7.2 Materials

- Standard molecular biology equipment and reagents for the following objectives:
  - Performing PCR reactions
  - Digesting and gel purifying DNA
  - Ligating and transforming DNA
  - Performing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and subsequent immunoblotting
- An appropriate vector system for expression of scFvs fused to an oligomerization domain (see Table 7.1 and Fig. 7.1, see also Chaps. 3 and 27)
- Cell disrupting instrument like a French Press (Aminco Rochester, NY, USA) or a TS 1.1 benchtop (Constant Systems Ltd. UK)

7.3 Procedure

The cloning of the miniantibodies (Figs. 7.1 and 7.2) follows standard procedures. We describe here the expression of miniantibodies in the periplasm of *E. coli* in shake flasks. A purification scheme using rapid coupled two-column purification is given in Chap. 27.

The procedure essentially follows the protocol as described earlier in Lindner and Plückthun (2001), which was based on Pack and Plückthun (1992). Detailed information on high-cell-density fermentation of miniantibodies on gram scales is given in Pack et al. (1993), Horn et al. (1996), Plückthun et al. (1996), and Schroeckh et al. (1996).

1. Select one of the presented formats for the chosen scFv (Table 7.1) and pick the appropriate vector (Table 7.2 and Fig. 7.2).

   Note: The vectors shown here and those in Chaps. 3 and 27 are modular and largely compatible because of matching restriction sites. In the pAK vectors, the scFv fragment is cloned between the (asymmetric) *Sfi*I sites, as discussed in detail in Chap. 3. The different dimerization or multimerization elements (Table 7.2) can be exchanged between the *Eco*RI and *Hin*dIII sites in most vectors (except, currently, the ones for the scFv-ZIPc and di-bi constructs). The rest of the backbone (e.g. between *Hin*dIII und *Xba*I) can be exchanged to coexpress a molecular chaperone (Chap. 27). In the pAK vectors, the region upstream of the scFv fragment (e.g. between *Xba*I and *Sfi*I, Fig. 7.2) can be exchanged to replace the Shine–Dalgarno sequence with a stronger version (Chap. 3).
2. If recloning from another expression vector without compatible *Sfi*I sites, design suitable primers, using, if desired, the information in Chap. 3 as a guide.

   Note: A protocol on how to PCR amplify an antibody with an unknown sequence from hybridoma or spleen cells and how to convert it into a scFv format compatible with this vector is given in Chap. 3.

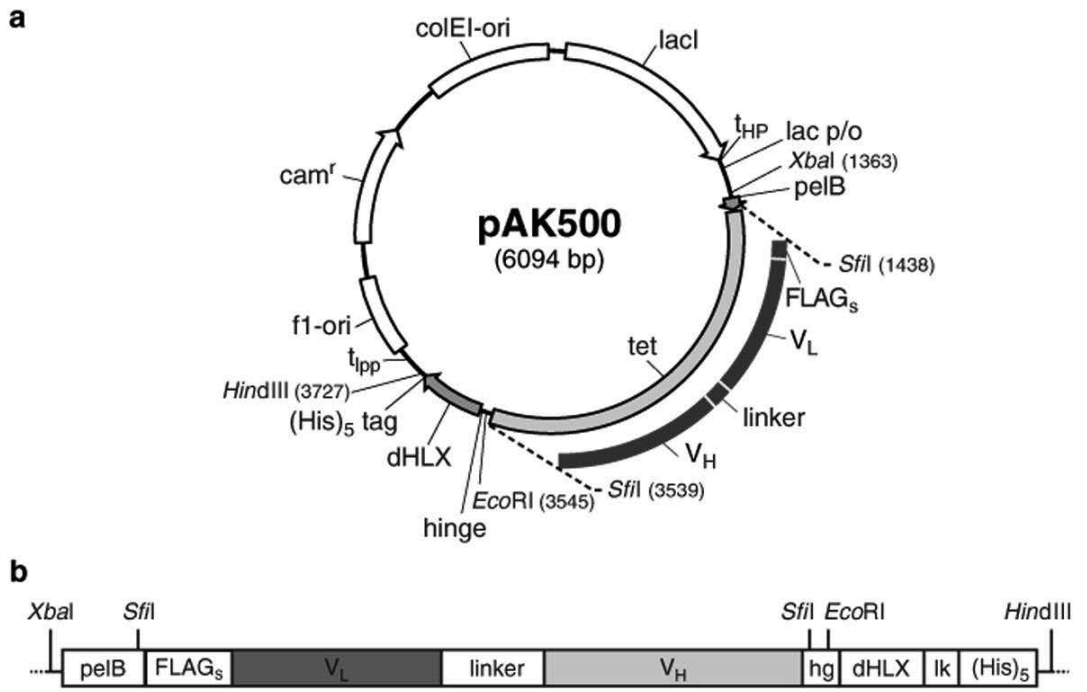


Fig. 7.2 *Expression vector pAK500 and schematic organization of the scFv-dHLX.* (**a**) Vector pAK500 (Krebber et al. 1997) encoding the dHLX cassette for the creation of a scFv construct in the scFv-dHLX format is shown as an example of a dimerization module. Other dimerization modules can be placed similarly between *Eco*RI and *Hin*dIII. The vector shown still contains the tetracycline-resistance cassette as stuffer to be replaced by the antibody scFv gene using the *Sfi*I cleavage sites (see Chap. 3). The resulting scFv insert is shown as the outer segment. Because of the compatibility between pAK vectors and the pJB series (Chaps. 3 and 27), elements (e.g. a stronger Shine Dalgarno sequence, absence of f1-ori, or a chaperone coexpression element) can be exchanged between different vectors. *lacI*: *lac* repressor; t_{HP}: strong upstream terminator to prevent read-through from LacI expression; *lac* p/o: *lac* promoter/operator; *pelB*: signal sequence (pectate lyase gene of *Erwinia carotovora*), modified to contain an *Sfi*I site; *tetR*: tetracycline resistance "stuffer" cassette (contains *tetA* and *tetR*-genes; 2,101 bp); *hinge*: murine IgG3 hinge region (see Table 7.2); *dHLX*: double helix element (see Table 7.2); $(His)_5$ *tag*: stretch of 5 histidine residues for IMAC purification (Lindner et al. 1992) and detection with an anti-his tag antibody (e.g. 3D5-phosphatase fusion (Lindner et al. 1997; Kaufmann et al. 2002)); t_{lpp}: downstream terminator; *f1 ori*: intergenic region of phage f1 (for production of single-stranded DNA); *cam^r*: chloramphenicol-acetyl-transferase gene; *colE1-ori*: plasmid replication origin (derived from pUC-plasmid series). (**b**) Schematic overview of the miniantibody construct (V_L-linker-V_H-dHLX fusion). *FLAG$_s$*: shortened (DYKD) version of FLAG tag for western blot detection (Knappik and Plückthun 1994); *hg*: murine IgG3 hinge region (see Table 7.2); *lk*: linker. Note: The size of the genetic elements is not drawn to scale. For more detail of the upstream and downstream region of the constructs, see Chap. 3

3. PCR amplify and clone the scFv in the selected vector. Confirm the correct arrangement of scFv and oligomerization domain in the final vector by DNA sequencing.

   Note: Some scFv fragments, especially when multimerized by any method, can become aggregation prone and potentially lead to growth defects of the

strain. In order to properly characterize the final construct before expression, it is recommended to add 1% glucose to all growth media in order to reduce expression before induction. The *lac* promoter/operator systems used here are under the control of the catabolite activator protein (CAP) and thus require the absence of glucose for full induction, or, conversely, are repressed by high glucose.

4. For expression, transform an *E. coli* host suitable for periplasmic expression (e.g. JM83 (Yanisch-Perron et al. 1985), RV308 (Maurer et al. 1980) or SB536 (Bass et al. 1996)).

   Note: JM83 is a generally robust strain that appears to lead to less lysis of the outer membrane upon periplasmic expression of some antibody fragments than some other strains (see also Chap. 27), RV308 is a strain that produces very little (inhibitory) acetate during growth to high cell densities and thus supports fermentation very well, and SB536 is deficient in two periplasmic proteases, HhoA (or DegQ) and HhoB (or DegS).

5. Inoculate a 20 ml pre-culture in LB medium (containing the appropriate antibiotic and 1% glucose) with a single bacterial colony harboring the plasmid encoding the respective scFv fragment. For this volume, use at least a 250 ml shaking flask. Incubate at 24°C overnight.

6. From this overnight culture, inoculate the main culture in SB medium containing 0.1% glucose at a starting OD_{600} of 0.1. Use a baffled shake flask for higher final cell densities to secure aeration. Shake at 24°C and add 1 mM IPTG (final concentration) at an OD_{600} of 0.5.

   Note: For most scFv fragments or miniantibody constructs, usage of only 0.1% glucose in the expression culture upon starting is recommended. In the majority of cases, this amount of glucose is enough to efficiently repress protein expression for 3–4 h until the culture has reached the OD required for induction. If higher concentrations of glucose are used, IPTG-induced protein expression might fail or be delayed in these CAP-regulated systems. However, there are some aggregation-prone scFv fragments which require the presence of 1% glucose at the time of inoculation. Whether this applies to the scFv of interest has to be tested individually.

   Note: The growth at room temperature is generally very beneficial for increasing the yield. At higher temperatures, not only does a more significant portion of many antibody fragments end up in the insoluble periplasmic fraction, but also incorrectly folded antibody fragments (or aggregates) interfere with membrane assembly, leading to an induced leakiness of the outer membrane and product loss.

   Note: If the expression vector carries the *skp* or *fkpA* gene (Chap. 27), much higher cell densities can be obtained, as the cells usually do neither lyse nor stop growing after induction (for details, see Chap. 27). Alternatively, a set of chaperones can be coexpressed on a second plasmid (Chap. 27).

7. Harvest the cells 4 h after induction by centrifugation (5,000 *g* for 10 min at 4°C)

   Note: This expression time is an average value, which depends on the aggregation properties of the construct and any proteolytic degradation, e.g.

in linker regions of fusion proteins. Robust constructs or constructs expressed in combination with overexpression of chaperones (Chap. 27) can be expressed for longer times.

8. Resuspend the cell pellet carefully in 1/100 column volume of loading buffer and add Benzonase (Merck) to a final concentration of 10 U/ml for removal of nucleic acids. Which loading buffer to use depends on the subsequent purification method chosen for the miniantibody. If the construct carries a His tag for IMAC purification (as for some of the constructs in Table 7.1 or pAK500, Fig. 7.2), it is recommended to use cold 50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0.

   Note: This and all subsequent steps should be carried out at 4°C in order to minimize protease activity and to stabilize the protein of interest.

   Note: To reduce protein degradation, protease inhibitors can be added to the solubilized cells. Proteolysis can be an issue for some fusion proteins, especially with positively charged residues in or near the linker region. However, the commercial protease inhibitor cocktails are mostly targeting eukaryotic proteases and are thus not very effective against *E. coli* proteases. Also, proteolysis, if it occurs by periplasmic enzymes, frequently begins during the induction phase, and can therefore only partially be combated with inhibitors.

   Note: In the product literature, Tris buffers are generally not recommended for IMAC, as their amines might interact with immobilized metal ions. However, we and others found that such buffer conditions do not influence the absorption of proteins containing hexa-histidine-tags, but rather keep some *E. coli* proteins from nonspecifically interacting with the chelating column matrix.

   Note: For troubleshooting, aliquots of the original culture and the supernatant after centrifugation should be kept and analyzed for scFv expression by SDS-PAGE and immunoblotting. These samples could pinpoint problems of the expression itself, compared to difficulties with the isolation and purification steps afterwards.

9. Disrupt the cells using a French Press (20,000 psi, 4°C in a cold room), the TS 1.1 benchtop, or sonification. For the French Press, perform at least three passages for optimal lysis of the cells. For all methods, take care that the cell suspension is not heated by the treatment.

10. Centrifuge the crude extract in order to separate insoluble cell debris from soluble protein (20,000 *g*, 30 min at 4°C). Carefully separate supernatant from pellet.

    Note: The soluble/insoluble distribution of the miniantibody expression can be analyzed by performing a western blot (see also Chap. 27). Since antibody fragments can form soluble aggregates, however, a mere inspection of western blots may be misleading. Therefore, a serious characterization of an antibody construct must include gel chromatography, ideally coupled with multi-angle light scattering. This will give a very clear description of the amount of soluble aggregates in a preparation, or their development over time.

Note: Successful transport to the periplasm can be inferred by the correct processing of the signal sequence. This can be detected by the anti-FLAG M1 antibody Sigma-Aldrich recognizing the processed FLAG tag at the very N-terminus ($^{+}H_3N$-DYKD...) (Knappik and Plückthun 1994), as the antibody does not recognize the tag when it is not at the N-terminus. This N-terminal short FLAG is present in the vector systems used here (this chapter, Chap. 3, Chap. 27).

11. Filter the supernatant through a 0.22 µm filter (use filters with low protein binding properties, e.g., Durapore filters from Millipore).
12. Apply the filtered supernatant of step 10 to the appropriate chromatography column.

    Note: Purification of antibody fragments using a rapid, directly coupled two-column procedure (IMAC and ion exchange chromatography) is presented in detail in Chap. 27.

7.4 Troubleshooting

While in general, the miniantibody strategy has been found to be quite robust, an intrinsic aggregation tendency of the scFv fragment is amplified by having several copies in one molecular assembly. If the protein of interest is mainly insoluble, the following procedures might be beneficial:

(a) Co-express one or several molecular chaperones which may increase the level of soluble expression (see Chap. 27, Bothmann and Plückthun 1998, 2000). Note, however, that the coexpression may merely shift insoluble aggregates to soluble aggregates. It is mandatory, therefore, to properly analyze the purified protein for oligomeric state by gel filtration, ideally coupled with multi-angle light scattering.

(b) Refold the protein from inclusion bodies. For this purpose, first reclone the scFv without any signal sequence into a plasmid with the strong T7-expression system (Ge et al. 1995). Refolding has to be optimized for each protein individually, but Huston et al. (1991), Ge et al. (1995) and Rudolph and Lilie (1996) give some initial guidelines. Commercial refolding kits are available, facilitating the screening for optimal conditions (Hampton Research, Laguna Niguel, CA, USA).

(c) Either introduce mutations in the scFv gene which may support proper folding or transplant the CDRs to a well-folding framework, thus leading to reduced aggregation. For an initial guidance, see Ewert et al. (2004). For additional discussions on this topic, see Knappik and Plückthun (1995), Jung and Plückthun (1997), Nieba et al. (1997), Willuda et al. (1999), Kaufmann et al. (2002), Honegger et al. (2009) or Kügler et al. (2009).

Acknowledgements This chapter is based on the original work of Peter Pack, Jörg Willuda and Susanne Kubetzko, with subsequent contributions from Kerstin Blank and Barbara Klinger.

References

Arndt KM, Pelletier JN, Müller KM, Alber T, Michnick SW, Plückthun A (2000) A heterodimeric coiled-coil peptide pair selected in vivo from a designed library-versus-library ensemble. J Mol Biol 295:627–639

Arndt KM, Müller KM, Plückthun A (2001) Helix-stabilized Fv (hsFv) antibody fragments: substituting the constant domains of a Fab fragment for a heterodimeric coiled-coil domain. J Mol Biol 312:221–228

Bass S, Gu Q, Christen A (1996) Multicopy suppressors of prc mutant *Escherichia coli* include two HtrA (DegP) protease homologs (HhoAB), DksA, and a truncated R1pA. J Bacteriol 178:1154–1161

Bothmann H, Plückthun A (1998) Selection for a periplasmic factor improving phage display and functional periplasmic expression. Nat Biotechnol 16:376–380

Bothmann H, Plückthun A (2000) The periplasmic *Escherichia coli* peptidylprolyl *cis*, *trans*-isomerase FkpA: I. Increased functional expression of antibody fragments with and without *cis*-prolines. J Biol Chem 275:17100–17105

Crothers DM, Metzger H (1972) The influence of polyvalency on the binding properties of antibodies. Immunochemistry 9:341–357

Deyev SM, Waibel R, Lebedenko EN, Schubiger AP, Plückthun A (2003) Design of multivalent complexes using the barnase*barstar module. Nat Biotechnol 21:1486–1492

Dürr E, Jelesarov I, Bosshard HR (1999) Extremely fast folding of a very stable leucine zipper with a strengthened hydrophobic core and lacking electrostatic interactions between helices. Biochemistry 38:870–880

Eisenberg D, Wilcox W, Eshita SM, Pryciak PM, Ho SP, DeGrado WF (1986) The design, synthesis, and crystallization of an alpha-helical peptide. Proteins 1:16–22

Ewert S, Honegger A, Plückthun A (2004) Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering. Methods 34:184–199

Ge L, Knappik A, Pack P, Freund C, Plückthun A (1995) Expressing antibodies in *Escherichia coli*. In: Borrebaeck C (ed) Antibody engineering, 2nd ed. Oxford University Press, London, pp 229–236

Harbury PB, Zhang T, Kim PS, Alber T (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. Science 262:1401–1407

Hill RB, Degrado WF (1998) Solution structure of alpha-2D, a native-like de novo designed protein. J Am Chem Soc 120:1138–1145

Holliger P, Prospero T, Winter G (1993) "Diabodies": small bivalent and bispecific antibody fragments. Proc Natl Acad Sci USA 90:6444–6448

Honegger A, Malebranche AD, Röthlisberger D, Plückthun A (2009) The influence of the framework core residues on the biophysical properties of immunoglobulin heavy chain variable domains. Protein Eng Des Sel 22:121–134

Horn U, Strittmatter W, Krebber A, Knüpfer U, Kujau M, Wenderoth R, Müller K, Matzku S, Plückthun A, Riesenberg D (1996) High volumetric yields of functional dimeric miniantibodies in *Escherichia coli*, using an optimized expression vector and high-cell-density fermentation under non-limited growth conditions. Appl Microbiol Biotechnol 46:524–532

Hu S, Shively L, Raubitschek A, Sherman M, Williams LE, Wong JY, Shively JE, Wu AM (1996) Minibody: A novel engineered anti-carcinoembryonic antigen antibody fragment (single-chain Fv-CH3) which exhibits rapid, high-level targeting of xenografts. Cancer Res 56:3055–3061

Huston JS, Mudgett-Hunter M, Tai MS, McCartney J, Warren F, Haber E, Oppermann H (1991) Protein engineering of single-chain Fv analogs and fusion proteins. Methods Enzymol 203: 46–88

Jeffrey PD, Gorina S, Pavletich NP (1995) Crystal structure of the tetramerization domain of the p53 tumor suppressor at 1.7 Ångstroms. Science 267:1498–1502

Jung S, Plückthun A (1997) Improving in vivo folding and stability of a single-chain Fv antibody fragment by loop grafting. Protein Eng 10:959–966

Kaufmann M, Lindner P, Honegger A, Blank K, Tschopp M, Capitani G, Plückthun A, Grütter MG (2002) Crystal structure of the anti-His tag antibody 3D5 single-chain fragment complexed to its antigen. J Mol Biol 318:135–147

Kellner C, Bruenke J, Stieglmaier J, Schwemmlein M, Schwenkert M, Singer H, Mentz K, Peipp M, Lang P, Oduncu F, Stockmeyer B, Fey GH (2008) A novel CD19-directed recombinant bispecific antibody derivative with enhanced immune effector functions for human leukemic cells. J Immunother 31:871–884

Kipriyanov SM, Moldenhauer G, Schuhmacher J, Cochlovius B, Von der Lieth CW, Matys ER, Little M (1999) Bispecific tandem diabody for tumor therapy with improved antigen binding and pharmacokinetics. J Mol Biol 293:41–56

Knappik A, Plückthun A (1994) An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments. Biotechniques 17:754–761

Knappik A, Plückthun A (1995) Engineered turns of a recombinant antibody improve its in vivo folding. Protein Eng 8:81–89

Krebber A, Bornhauser S, Burmester J, Honegger A, Willuda J, Bosshard HR, Plückthun A (1997) Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. J Immunol Methods 201:35–55

Kubetzko S, Balic E, Waibel R, Zangemeister-Wittke U, Plückthun A (2006) PEGylation and multimerization of the anti-p185HER-2 single chain Fv fragment 4D5: effects on tumor targeting. J Biol Chem 281:35186–35201

Kügler M, Stein C, Schwenkert M, Saul D, Vockentanz L, Huber T, Wetzel SK, Scholz O, Plückthun A, Honegger A, Fey GH (2009) Stabilization and humanization of a single-chain Fv antibody fragment specific for human lymphocyte antigen CD19 by designed point mutations and CDR-grafting onto a human framework. Protein Eng Des Sel 22:135–147

Lindner P, Plückthun A (2001) Miniantibodies. In: Kontermann R, Dübel S (eds) Antibody engineering. Springer, Berlin, pp 637–647

Lindner P, Guth B, Wülfing C, Krebber C, Steipe B, Müller F, Plückthun A (1992) Purification of native proteins from the cytoplasm and periplasm of *Escherichia coli* using IMAC and histidine tails: a comparison of proteins and protocols. Methods 4:41–56

Lindner P, Bauer K, Krebber A, Nieba L, Kremmer E, Krebber C, Honegger A, Klinger B, Mocikat R, Plückthun A (1997) Specific detection of his-tagged proteins with recombinant anti-His tag scFv-phosphatase or scFv-phage fusions. Biotechniques 22:140–149

Maurer R, Meyer B, Ptashne M (1980) Gene regulation at the right operator (OR) bacteriophage lambda. I. OR3 and autogenous negative control by repressor. J Mol Biol 139:147–161

Mittl PR, Chene P, Grütter MG (1998) Crystallization and structure solution of p53 (residues 326–356) by molecular replacement using an NMR model as template. Acta Crystallogr D Biol Crystallogr 54:86–89

Müller KM, Arndt KM, Plückthun A (1998a) A dimeric bispecific miniantibody combines two specificities with avidity. FEBS Lett 432:45–49

Müller KM, Arndt KM, Plückthun A (1998b) Model and simulation of multivalent binding to fixed ligands. Anal Biochem 261:149–158

Müller KM, Arndt KM, Strittmatter W, Plückthun A (1998c) The first constant domain (CH1 and CL) of an antibody used as heterodimerization domain for bispecific miniantibodies. FEBS Lett 422:259–264

Nieba L, Honegger A, Krebber C, Plückthun A (1997) Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. Protein Eng 10:435–444

O'Shea EK, Klemm JD, Kim PS, Alber T (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 254:539–544

Pack P, Plückthun A (1992) Miniantibodies: use of amphipathic helices to produce functional, flexibly linked dimeric Fv fragments with high avidity in *Escherichia coli*. Biochemistry 31:1579–1584

Pack P, Kujau M, Schroeckh V, Knüpfer U, Wenderoth R, Riesenberg D, Plückthun A (1993) Improved bivalent miniantibodies, with identical avidity as whole antibodies, produced by high cell density fermentation of *Escherichia coli*. Biotechnology (NY) 11:1271–1277

Pack P, Müller K, Zahn R, Plückthun A (1995) Tetravalent miniantibodies with high avidity assembling in *Escherichia coli*. J Mol Biol 246:28–34

Plückthun A, Pack P (1997) New protein engineering approaches to multivalent and bispecific antibody fragments. Immunotechnology 3:83–105

Plückthun A, Krebber A, Krebber C, Horn U, Knüpfer U, Wenderoth R, Nieba L, Proba K, Riesenberg D (1996) Producing antibodies in *Escherichia coli*: Fom PCR to fermentation. In: McCafferty J, Hoogenboom H (eds) Antibody engineering: a practical approach. IRL Press, Oxford, pp 203–252

Rheinnecker M, Hardt C, Ilag LL, Kufer P, Gruber R, Hoess A, Lupas A, Rottenberger C, Plückthun A, Pack P (1996) Multivalent antibody fragments with high functional affinity for a tumor-associated carbohydrate antigen. J Immunol 157:2989–2997

Rudolph R, Lilie H (1996) In vitro folding of inclusion body proteins. FASEB J 10:49–56

Schroeckh V, Kujau M, Knüpfer U, Wenderoth R, Mörbe J, Riesenberg D (1996) Formation of recombinant proteins in *Escherichia coli* under control of a nitrogen regulated promoter at low and high cell densities. J Biotechnol 49:45–58

Todorovska A, Roovers RC, Dolezal O, Kortt AA, Hoogenboom HR, Hudson PJ (2001) Design and application of diabodies, triabodies and tetrabodies for cancer targeting. J Immunol Methods 248:47–66

Willuda J, Honegger A, Waibel R, Schubiger PA, Stahel R, Zangemeister-Wittke U, Pluckthun A (1999) High thermal stability is essential for tumor targeting of antibody fragments: engineering of a humanized anti-epithelial glycoprotein-2 (epithelial cell adhesion molecule) single-chain Fv fragment. Cancer Res 59:5758–5767

Willuda J, Kubetzko S, Waibel R, Schubiger PA, Zangemeister-Wittke U, Plückthun A (2001) Tumor targeting of mono-, di-, and tetravalent anti-p185(HER-2) miniantibodies multimerized by self-associating peptides. J Biol Chem 276:14385–14392

Woolfson DN (2005) The design of coiled-coil structures and assemblies. Adv Protein Chem 70:79–112

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119

Zhang J, Tanha J, Hirama T, Khieu NH, To R, Tong-Sevinc H, Stone E, Brisson JR, MacKenzie CR (2004) Pentamerization of single-domain antibodies from phage libraries: a novel strategy for the rapid generation of high-avidity antibody reagents. J Mol Biol 335:49–56

3.5 Improving Expression of scFv Fragments by Co-expression of Periplasmic Chaperones

Jonas V. Schaefer and Andreas Plückthun

in **Antibody Engineering** (Kontermann, R., and Dübel, S., eds) Vol. 2, 2nd edit., 2010,pp. 345-361, Springer Verlag, Berlin Heidelberg, Germany

Chapter 27
Improving Expression of scFv Fragments by Co-expression of Periplasmic Chaperones

Jonas V. Schaefer and Andreas Plückthun

Abbreviations

IPTG	Isopropylthiogalactoside
PBS	Phosphate buffered saline
scFv	Single-chain Fv fragment
tet	Tetracycline
ELISA	Enzyme-linked Immunosorbent Assay
LB	Luria–Bertani media
SB	Super broth media

27.1 Introduction

For more than 20 years now, periplasmic expression in *Escherichia coli* has become the standard technology for preparing functional antibody fragments in a rapid and convenient way (Skerra and Plückthun 1988; Plückthun et al. 1996). The criteria of choosing either the Fab or single-chain Fv fragment (scFv) format, the properties of suitable expression vectors, as well as the influence of the *E. coli* strain used have been extensively summarized elsewhere (Plückthun et al. 1996). However, even when considering all these components and experimental conditions, the yield of recombinant antibody fragments is still highly variable, mainly being a direct consequence of the primary sequence and its sequence-dependent propensity to lead to aggregation-prone folding intermediates. In general, periplasmic folding is the yield-limiting step, being strongly influenced by the amino acid composition of the antibody to be expressed (Wörn and Plückthun 2001; Ewert et al. 2004). However, yield is not the only property influenced as the protein sequence also

J.V. Schaefer and A. Plückthun (✉)
Biochemisches Institut, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
e-mail: plueckthun@bioc.uzh.ch

determines stability and resistance against aggregation upon storage of the purified protein. Since these properties cannot be changed by expression conditions, antibody sequence alteration must be seen in conjunction with choosing an appropriate expression system, and this includes chaperone co-expression.

Two principal methods have proven to be successful for improving antibody sequences: a "rational" approach and a "directed evolution" one. The rational approach is based on alignments of the particular antibody sequence to that of well-expressing fragments (Knappik and Plückthun 1995; Wörn et al. 2000; Ewert et al. 2003, 2004; Honegger et al. 2009), an analysis of exposed hydrophobic residues (Nieba et al. 1997), or the grafting of CDRs onto a stable and well-folding framework (Jung and Plückthun 1997; Willuda et al. 1999; Kügler et al. 2009). In a directed evolution approach, the protein is subjected to an evolutionary pressure, which rewards stability and expression (Jung et al. 1999; Jermutus et al. 2001; Schimmele and Plückthun 2005). When starting from a given antibody, such protein engineering constitutes — undeniably — a significant effort. The ability to rapidly characterize the given antibody fragment will be critical whenever a choice between various fragments with different binding properties has to be made. For this purpose, significant amounts of properly folded protein are necessary.

Therefore, we discuss here the co-expression of periplasmic factors improving the yield of soluble and correctly folded antibody. Because of their conserved intradomain disulfide bonds, antibody fragments need to be secreted to an oxidizing compartment for correct folding (Skerra and Plückthun 1988), this being the periplasmic space in bacteria. While some antibodies have been engineered to fold in the absence of disulfides (Proba et al. 1998) and others have been expressed (Proba et al. 1995; Levy et al. 2001) in *E. coli* mutant strains with altered cytoplasmic redox machinery where cytoplasmic disulfides can accumulate to some extent (Ortenberg and Beckwith 2003), this chapter deals with periplasmic expression.

The effect of overexpressing molecular chaperones and other folding modulators on the yield of foreign proteins has been reviewed (Wall and Plückthun 1995; Kolaj et al. 2009). Since the folding of the antibody takes place *after* its secretion, periplasmic factors are of greatest interest in this regard. Nonetheless, the overexpression of cytoplasmic factors has also been attempted in the hope of improving yield of soluble antibody (Söderlind et al. 1995; Hu et al. 2007).

In the bacterial periplasm, three types of folding modulators have been identified that may play a role with the folding of exogenous proteins in *E. coli*: (1) the disulfide-bond-forming (Dsb) machinery (Kadokura et al. 2003; Ortenberg and Beckwith 2003), with the periplasmic proteins DsbA and DsbC, and to some degree the specialized DsbE and DsbG (DsbB and DsbD being transmembrane proteins for regenerating the periplasmic factors); (2) the four periplasmic proteins with peptidyl prolyl *cis/trans* isomerase (PPI) activity (Galat 2003), PpiA (RotA), PpiD, FkpA, and SurA; and (3) the protein Skp with chaperone activity (see below), and finally the protease DegP (Skorko-Glonek et al. 2008) suspected to also have chaperone activity at low temperature.

There is clear evidence that the spectrum of activities of these proteins is overlapping. The dimeric peptidyl prolyl *cis/trans* isomerase FkpA has chaperone

activity, most clearly shown by its improvement of the periplasmic expression of scFv fragments that do not even have a *cis* proline (Bothmann and Plückthun 2000; Ramm and Plückthun 2000). The dimeric DsbC, while showing disulfide isomerase activity, is also thought to have chaperone activity (Chen et al. 1999; Zhao et al. 2003; Segatori et al. 2004) and has been observed to help against periplasmic lysis. However, clear evidence is lacking that an increased peptidyl prolyl *cis/trans* isomerase activity and an increased disulfide isomerization activity are actually *per se* beneficial for antibody scFv fragments, as opposed to the observed favorable effects being entirely due to the built-in chaperone activities of FkpA and DsbC, and possibly other factors (Bothmann and Plückthun 2000; Ramm and Plückthun 2000; Sandee et al. 2005). These enzymatic activities may become of great importance in other antibody constructs, however, e.g., those with additional disulfide bonds.

We have taken two approaches to tackle the problem of soluble expression of scFvs. First, we have previously identified factors that increase the functional expression of antibody constructs by a selection approach and have designed appropriate co-expression vectors. Second, we have created a modular system that allows a flexible co-expression of many factors with virtually any antibody expression vector.

For identification of the crucial factors for antibody expression, we used a phage display system, which displayed a constant, poorly folded antibody fragment and a library of co-expressed genes (Bothmann and Plückthun 1998; Bothmann and Plückthun 2000). With this enrichment strategy, we identified two periplasmic factors with beneficial, chaperone-like properties, both increasing the folding efficiency of scFvs and, consequently, their yield in the periplasm. The first factor identified, Skp (for 17 kDa protein), is a basic periplasmic protein that has been found to specifically interact with outer membrane proteins, assisting their transport through the periplasm (De Cock et al. 1999; Schäfer et al. 1999), and it may similarly interact with the folding antibody. It has recently been found to interact also with some periplasmic *E. coli* proteins (Jarchow et al. 2008). The second factor, FkpA, is a periplasmic peptidyl prolyl *cis/trans* isomerase, which also acts as a chaperone (Bothmann and Plückthun 2000; Ramm and Plückthun 2000), perhaps the more important property. The effects of Skp and FkpA (increasing the scFv yield by up to a factor 10) appear to be specific for every antibody variant, as neither additivity nor synergy was observed. However, we never noticed a negative influence of the co-expression of either Skp or FkpA onto the scFv level up to now — in some cases it had simply no effect, notably when the antibody did not show significant aggregation tendencies to begin with.

While Skp and FkpA have been experimentally identified by an enrichment strategy to be helpful for antibody fragments, a more generic co-expression strategy can be useful as well. Therefore, we designed a plasmid series, named pCH, for the overexpression of the thiol-disulfide oxidoreductases DsbA and DsbC, based on the pTUM4 vector (Schlapschy et al. 2006). The coding sequence for Skp, as well as FkpA and SurA (another peptidyl prolyl *cis/trans* isomerase with suspected chaperone activity, implicated in the delivery of proteins across the periplasm to

the outer membrane),was also included. The main reason for altering the existing pTUM4 plasmid was to create a new modular structure that should be compatible to virtually any antibody expression plasmid. This was achieved using a modular design previously utilized by Lutz and Bujard (1997), allowing a convenient exchange of both the origin of replication as well as the genes conferring antibiotic resistance by unique restriction sites. Thus, we created a set of plasmids, carrying different combinations of origin of replications (ColE1, p15A, and pSC101, each resulting in a different number of intracellular plasmid copies) in conjunction with the genes conferring resistance to ampicillin, kanamycin, chloramphenicol, or tetracycline, respectively. With this variety of origins and resistance genes, the pCH series is compatible with virtually all conventional antibody expression vectors. The choice of different origins allows one to control the level of chaperone co-expression based on different plasmid copy numbers. It also safeguards against plasmid incompatibility (which can lead to the loss of one of the plasmids), even though this may be less of a concern in high copy number plasmids (Velappan et al. 2007). Initial results indicate an overall yield increase of antibody fragments in a variety of formats upon co-transformation of suitable *E. coli* hosts with members of this vector series.

27.2 Materials

- Standard molecular biology equipment and reagents for
  - Isolating genomic DNA from *E. coli* (e.g., Qiagen DNeasy Blood & Tissue Kit),
  - Performing PCR reactions,
  - Cutting and gel-purifying DNA (e.g., Sigma-Aldrich GenElute Gel Extraction Kit),
  - Ligating and transforming DNA,
  - Conducting an enzyme-linked immunosorbent assay (ELISA), and
  - Performing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and subsequent immunoblotting;
- An appropriate expression system to produce histidine-tagged antibody fragments in the periplasm, such as the pAK system (see Chap. 3 or Krebber et al. 1997);
- Cell disrupting instrument like a French Press (Aminco Rochester, NY, USA) with 4 ml cell and 40 ml cell or a TS 1.1 benchtop (Constant Systems Ltd. UK);
- An automated LC-System: e.g., BioCAD workstation (e.g., PerSeptive Biosystems, acquired by Applied Biosystems) with dual-channel variable-wavelength UV/visible detector, semipreparative flow cell (Perkin Elmer), fraction collector Advantec SF-2120 (Toyo Roshi International), or equivalent system;
- POROS20 MC/M 4.6 mm/100 mm (metal chelate) (Applied Biosystems);
- POROS20 HQ/M 4.6 mm/100 mm (anion exchange) (Applied Biosystems);
- POROS20 HS/M 4.6 mm/100 mm (cation exchange) (Applied Biosystems);

- Imidazole stock solution (1 M) adjusted to pH 7 with acetic acid.
  Note: Make sure to adjust the pH of the imidazole stock solution using *acetic acid*, and not with HCl, in order to keep the ionic strength low (otherwise the protein might run through the coupled downstream ion exchange column);
- NaCl stock solution (3 M);
- $NiCl_2$ (200 mM);
- Distilled water.

PBS (PBST)
Na_2HPO_4 (10 mM);
KH_2PO_4 (1.8 mM);
KCl (2.7 mM);
NaCl (137 mM, pH 7.4);
for PBST, also add Tween 20 to a final concentration of 0.05%

Extraction Buffer
Sucrose, 20% (w/v);
EDTA (1 mM);
Tris-HCl (100 mM, pH 8.0).

Solubilization Buffer
Urea (2 M);
EDTA (1 mM);
Glycylglycine (10 mM, pH 7.5).

MHA Buffer (5× stock solution is given)
Mes (33 mM);
Hepes (33 mM);
Na-acetate (33 mM; adjust to pH 7.5 with NaOH unless a different pH is indicated below).

27.3 Procedure

27.3.1 Construction of Vectors for the Co-expression of Periplasmic Chaperones

Co-expression the chaperones mentioned above can either be driven from expression cassettes within the same vector or from separate plasmids used in co-transformations. In the following, the design and cloning of such vectors is described.

27.3.1.1 Cloning of scFv Fragments from pAK/pJB into Vectors Overexpressing Periplasmic Chaperones (pHB110, pHB610, pJB33)

1. Excise the expression cassette coding for the scFv antibody fragment from the relevant pAK/pJB vector (described in detail in Chap. 3) by digestion with *Xba*I and *Hin*dIII. Use 2 µg purified plasmid DNA and incubate at 37°C for 2 h in a

total volume of 50 µl containing 5 µl 10 × NEBuffer 2 (NEB), 5 µl 10 × BSA, and 20 units of each *Xba*I (NEB) and *Hin*dIII (NEB).

Note: Procedure 27.3.1.1 describes the co-expression of periplasmic chaperones on the expression vector itself. These vectors have compatible restriction sites with the phage display vectors described in Chap. 3. The vectors differ (Fig. 27.1) in whether they also allow phage display and thus have a moderately strong translation initiation region (pHB110, pHB610), or only allow periplasmic expression and have a strong translation initiation region (pJB33), and in whether they co-express Skp or FkpA. All vectors have compatible restriction sites.

2. Digest appropriate amounts of vector (pHB110, pHB610, or pJB33) with *Xba*I and *Hin*dIII (removing the *tet*-cassette, see Fig. 27.1) for 2 h at 37°C under the same conditions as above. Also, dephosphorylate the cut vector by adding calf intestinal alkaline phosphatase (CIP, NEB; 0.5 unit/µg vector) to the digestion after 1 h.

   Note: Dephosphorylation should not be necessary because of the asymmetric overhangs. However, we always include this step to eliminate any risk of re-ligation of single-cut vector.

3. Purify the digested scFv antibody genes and vector by preparative agarose gel electrophoresis in combination with the GenElute gel extraction kit (Sigma-Aldrich).

   Note: For pure preparations of a completely digested vector, it is very important not to overload the agarose gel. Furthermore, the gel electrophoresis has to be run long enough to separate small amounts of undigested or single-cut vector from the digested vector band.

4. Ligate 50 ng cut vector with the scFv expression cassette (molar ratio vector to insert 1:5) with 5 units T4 DNA ligase (NEB) in the presence of 1 × T4 DNA ligase buffer in 10 µl volume. Incubate for 2 h at room temperature or overnight at 16°C.

5. Transform 50 µl chemocompetent *E. coli* host cells suitable for periplasmic expression (e.g., JM83 (Yanisch-Perron et al. 1985), RV308 (Maurer et al. 1980), or SB536 (Bass et al. 1996)) with 5 µl of the ligation mix by heat-shock for 45 s at 42°C; add 500 µl of Luria–Bertani media (LB) media after 2 min incubation on ice and incubate for 60 min shaking at 37°C. Plate all transformed cells on LB, 1% glucose, chloramphenicol (30 µg/ml) agar plates and incubate overnight at 37°C.

   Note: JM83 is a generally robust strain that appears to lead to less lysis of the outer membrane upon periplasmic expression of some antibody fragments than some other strains. RV308 is a strain that produces very little (inhibitory) acetate during growth to high cell densities and thus supports fermentation very well. SB536 is deficient in two periplasmic proteases, HhoA (or DegQ) and HhoB (or DegS).

27.3.1.2 Cloning of *skp*/*fkpA* in Other Expression Vectors

Both the *skp* and *fkpA* genes can be conveniently obtained by digestion and purification from the vector pHB110 or pHB610, respectively (digested either

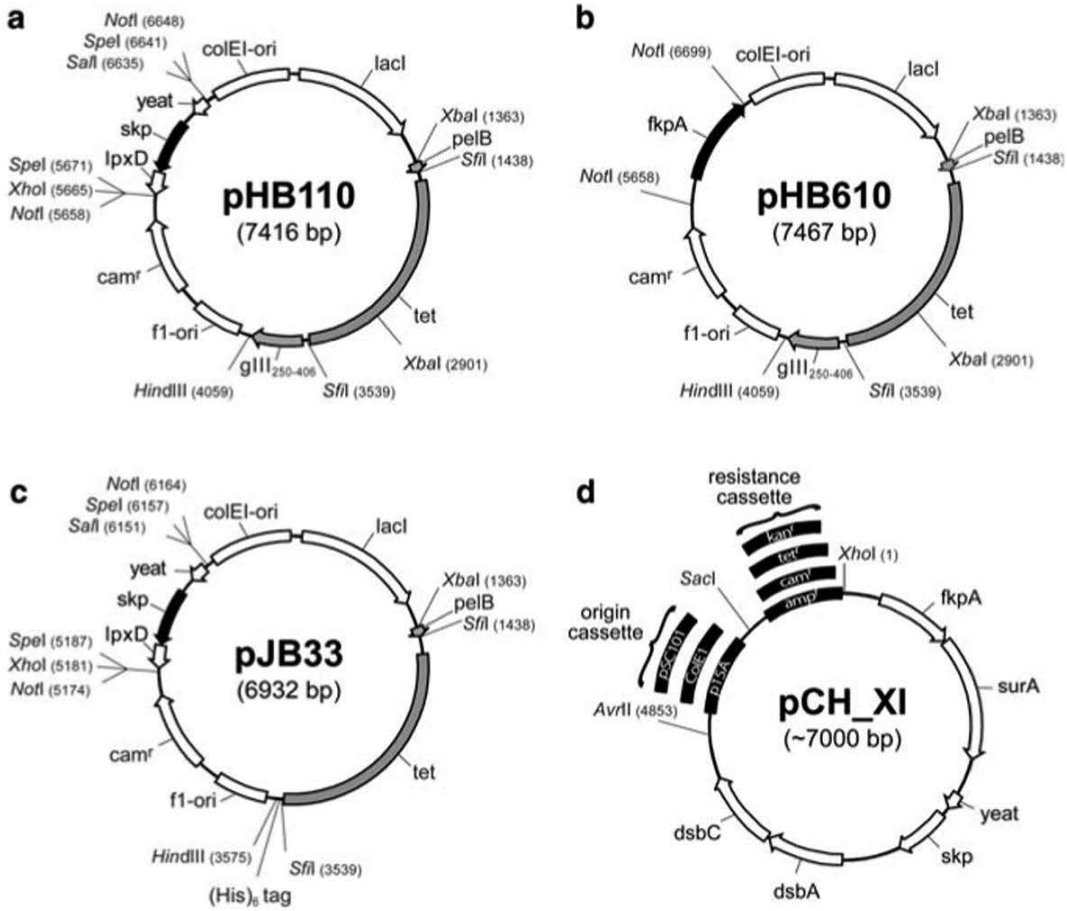


Fig. 27.1 *Vectors and cloning strategies*. The vectors pHB110, pHB610, and pJB33 all contain a chloramphenicol resistance gene (*cam*^{r}) as well as a tetracycline resistance "stuffer" cassette (*tet*, 2101 bp) (Krebber et al. 1997), which will be replaced by the antibody fragment (see Fig. 2.4). This stuffer is shown only schematically and contains the genes for *tetA* and *tetR* without making any fusion protein with upstream or downstream elements in the vector (for details, see Chap. 3). Vectors pHB110 and pHB610 allow either phage display (upon introducing an scFv cassette without stop codon, resulting in a fusion with the phage gene III) or periplasmic expression (if a stop codon is present at the end of the scFv gene), leading to moderate translation levels. In contrast, vector pJB33 leads to an enhanced periplasmic expression (due to the strong Shine–Dalgarno sequence SDT7g10 from T7 phage) and permits subsequent IMAC purification of the antibody fragment (see Chap. 3). Because of their compatible design, elements (e.g., the strong Shine– Dalgarno sequence) can be exchanged between vectors. (**a**) Vector pHB110 containing the *skp* cassette with flanking genes (in the form it was enriched during panning (Bothmann and Plückthun 1998)). This vector can also be used as a source of *skp* after digestion with *Not*I, *Spe*I, and *Sal*I/*Xho*I. (**b**) Vector pHB610 containing *fkpA*, excisable using *Not*I. (**c**) Vector pJB33 with stronger translation initiation region for high yield expression of scFv (see Chap. 3). (**d**) Schematic overview of the pCH series, encoding five different chaperones. As indicated, both the cassette for the origin of replication as well as that for the antibiotic resistance is exchangeable using *Avr*II/*Sac*I or *Sac*I/*Xho*I, respectively. *lpxD*: the first 65 aa of UDP-3-O-[hydroxymyristoyl]-glucosamine-*N*-acyltransferase, *yeaT*: the last 49 aa of YeaT (outer membrane proteins involved in the insertion of other outer membrane proteins). The sequence of the vectors is available from the authors upon request

skp-for:	5'	NNN	NNN	XXX	XXX	GAT	CCA	AGC	AAT	ATC	CGT	ATG	TCT	GC		3'
skp-rev:	5'	NNN	NNN	XXX	XXX	TTA	TTT	AAC	CTG	TTT	CAG	TAC	GTC	GGC		3'
fkpA-for:	5'	NNN	NNN	XXX	XXX	GAT	TCA	CCT	CTT	TTG	TCG	AAT	GGT	C		3'
fkpA-rev:	5'	NNN	NNN	XXX	XXX	TTA	TTT	TTT	AGC	AGA	ATC	TGC	GGC	3		3'

Fig. 27.2 *Primers used for the amplification of skp and fkpA*. XXX XXX stands for the restriction site used for subcloning, while NNN NNN represents the additional bases flanking the restriction sites necessary for efficient cleavage (see e.g., the New England Biolabs catalog)

with *Not*I or alternatively with *Spe*I and *Xho*I/*Sal*I; see Fig. 27.1). By PCR amplification, new restriction sites can, of course, be added to insert them into any desired expression vector. These amplified fragments might thus also be useful for insertion into vectors used for expressing antibody fragments other than scFvs, and other periplasmic protein altogether (Fig. 27.1).

As neither for *skp* nor for *fkpA* the exact limits of their promoters have been experimentally verified, we recommend using the PCR primers specified in Fig. 27.2 for amplifying the genes (if they are to be expressed under their own promoters) from the vectors pHB110 or pHB610 or, alternatively, from genomic *E. coli* DNA:

1. Isolate genomic DNA from *E. coli* using the DNeasy Blood & Tissue Kit (Qiagen) as described by the manufacturer.
2. Perform PCR amplification of *skp* or *fkpA* with the above mentioned primers according to standard protocols.

   Note: We recommend using high-fidelity, proofreading polymerases (e.g., Phusion High-Fidelity DNA Polymerase from Finnzymes). At the beginning of the PCR reaction, the annealing temperatures for the above mentioned primers are set to the theoretical values of 56–59°C. However, we recommend increasing this temperature after the first five cycles (depending on the additional nucleotides added as overhang), as the amplified PCR product including this overhang will serve itself as template DNA for further amplification.
3. Digest the PCR product with the appropriate restriction enzymes, and ligate it into your favorite expression vector (also see steps 27.3.1.1.4–5).

27.3.1.3 Co-expression of Chaperones Encoded on a Second Plasmid

As an alternative to co-expressing the desired chaperones from the same plasmid, double transformation of *E. coli* hosts with two plasmids (one encoding the antibody fragment to be expressed and the second one harboring the genes for the chaperone(s)) is an option. However, the plasmids must possess different antibiotic resistance and preferentially different origin of replications, even though this is not strictly required when they are of high copy number (Velappan et al. 2007). The choice of different origins can be beneficial since chaperone expression can be tuned by differences in the plasmid copy number. Therefore, we used a modular design based on the pZ vector system developed by Lutz and Bujard (1997) for the pCH vector series leading to the constitutive overexpression of the chaperones

DsbA, DsbC, FkpA, SurA, and Skp. As mentioned above, this modular structure provides the chance to choose between the ColE1 origin of replication, (resulting in 50–70 intracellular plasmid copies), p15A origin (20–30 copies), and pSC101 origin (~10 copies), as well as cassettes encoding resistance to ampicillin, kanamycin, chloramphenicol, or tetracycline. These vectors should therefore be compatible with virtually any existing expression plasmid. For further details also refer to the legend of Fig. 27.1.

1. Transform suitable *E. coli* host cells with both plasmids, coding for the scFv and chaperones, respectively, as described in 27.3.1.1.5. For periplasmic expression, JM83 (Yanisch-Perron et al. 1985) is a robust host, but many other strains can be used (see 27.3.1.1.5).
2. Plate all transformed cells on LB, 1% glucose agar plates containing both appropriate antibiotics and incubate overnight at the desired temperature. It may be useful to test the effect of co-expression both at room temperature and at 37°C.

27.3.2 Small-Scale Expression of scFv Antibody Fragments

1. Inoculate 10 ml SB medium (per l, 35 g tryptone, 20 g yeast extract, 5 g NaCl, pH 7.5), containing the appropriate antibiotic(s) and 0.1% glucose, with a single colony of transformed *E. coli*, harboring the plasmid encoding the respective scFv fragment, and, if applicable, the plasmid co-expressing the chaperones. Grow the culture at 24°C and induce with 1 mM isopropylthiogalactoside (IPTG) (final concentration) at an OD_{600} of 0.5.

   Note: This procedure aims at analyzing the relative amounts of soluble protein for different constructs and/or chaperone co-expression.

   Note: The growth at room temperature is in general very beneficial for increasing the yield. At higher temperature, not only does a more significant portion of many antibody fragments end up in insoluble periplasmic fractions, but also incorrectly folded antibody fragments (or aggregates) interfere with membrane assembly, leading to an induced leakiness of the outer membrane and product loss.

   Note: Use only 0.1% glucose or less in the expression culture upon starting. This amount of glucose is enough to efficiently repress protein expression for 3–4 h until the culture has reached the OD required for induction. If higher concentrations of glucose are used, IPTG-induced protein expression might fail or be delayed.

   Note: When analyzing many constructs in parallel, it might be beneficial to grow overnight pre-cultures and inoculate the final expression culture at a starting OD_{600} of 0.1. This will lead to growth synchronization of the cultures and therefore synchronize the time points where the OD for IPTG induction has been reached.

2. Harvest the cells 4 h after induction by centrifugation (5,000 *g* for 10 min at 4°C).

   Note: This expression time is an average value, which depends on the aggregation properties of the construct and any proteolytic degradation, e.g., in linker regions of fusion proteins. Robust constructs can be expressed for longer times.

   Note: For troubleshooting, aliquots of the original culture and the supernatant after centrifugation should be kept and analyzed for scFv expression by SDS-PAGE and immunoblotting. These samples could pinpoint problems of the expression itself, compared to difficulties with the isolation and purification steps afterwards.
3. Resuspend the cells carefully in 0.5 ml pre-cooled extraction buffer on ice, and measure the OD_{600}. Do not lyse the bacteria. Add lysozyme (Sigma-Aldrich; 100 µg/ml) and incubate for 1 h on ice.

   Note: This procedure will destabilize both the *E. coli* peptidoglycan and the outer membrane, allowing soluble contents of the periplasm to leak out.
4. Centrifuge bacteria at 5,000 *g* for 10 min at 4°C and carefully transfer the supernatant (soluble periplasmic fraction) to a fresh Eppendorf tube.
5. Dissolve the pellet in 0.5 ml solubilization buffer (insoluble fraction).

   Note: This solubilization can be performed overnight, shaking at 4°C if necessary. This concentration of urea will in general be sufficient to dissolve periplasmic aggregates.
6. Normalize all fractions to the same OD_{600} of the original culture.

   Note: Make fractions comparable between cultures (correct for OD_{600}) and within a culture such that aliquots from the soluble and insoluble fractions can be compared easily.
7. For ELISA, coat suitable microtiter plates with the appropriate antigen overnight at 4°C according to standard protocols (see, e.g., Thorpe and Kerr 1994). Mix a defined amount of normalized soluble fraction with 2% skimmed milk in PBST and apply to the blocked ELISA plate. Subsequently, perform detection as, e.g., described in Thorpe and Kerr (1994).

   Note: If soluble antigen is available, include a competition ELISA control showing that free antigen is able to compete with bound antigen for binding to distinguish nonspecific "sticky" from specifically binding scFvs.
8. For western blot analysis, load defined amounts of soluble and insoluble protein fractions (also including samples taken in step 27.3.2.2 boiled in SDS-loading buffer to have a control of the total expression) on a 15% SDS-PAGE under reducing conditions. Perform standard immunoblotting according to the protocols described in Sambrook and Russell (2001).

   Note: To judge the effect of a construct or chaperone co-expression, it is important to evaluate both the total amount in the soluble fraction, as well as the ratio of soluble to insoluble protein.

   Note: The detected soluble protein may not necessarily be functional as it might consist of soluble aggregates (see next section).

   Note: Successful transport to the periplasm can be inferred from the correct processing of the signal sequence. This can be detected by the M1 antibody (Sigma-Aldrich) recognizing the processed FLAG tag at the very N-terminus

($^{+}H_3N$-DYKD...) (Knappik and Plückthun 1995), as the antibody does not recognize the tag when it is not at the N-terminus. This N-terminal short FLAG is present in the vector systems used here (this chapter, Chap. 3 and 7).

27.3.3 Large-Scale Expression

The single-chain Fv fragment carrying a C-terminal hexa-histidine tag (e.g., after expression from plasmid pJB33) can by purified by rapid two-column chromatography as described below (Sect. 27.3.3 and 27.3.4). This protocol is designed for 5–10 g wet weight of *E. coli* cells, corresponding to about 1 l of baffled shake-flask culture.

1. Inoculate a pre-culture of 10 ml SB medium, containing the appropriate antibiotic(s) and 1% glucose, with a single colony of *E. coli*, harboring the plasmid encoding the respective scFv fragment, and optionally a co-expression plasmid for chaperones. Incubate at 24°C overnight.
2. From this overnight culture, inoculate the main culture of 1 l SB medium containing 0.1% glucose at a starting OD_{600} of 0.1. Grow the culture at 24°C in a baffled shake flask for higher final cell densities and induce with 1 mM IPTG (final concentration) at an OD_{600} of 0.5.

   Note: Use only 0.1% glucose in the expression culture upon starting. This amount of glucose is enough to efficiently repress protein expression for 3–4 h until the culture has reached the OD required for induction. If higher concentrations of glucose are used, IPTG-induced protein expression might fail or be delayed.

   Note: The growth at room temperature is in general very beneficial for increasing the yield. At higher temperature, not only does a more significant portion of many antibody fragments end up in insoluble periplasmic fractions, but also incorrectly folded antibody fragments (or aggregates) interfere with membrane assembly, leading to an induced leakiness of the outer membrane and product loss.
3. Harvest the cells ca. 4 h after induction by centrifugation (5,000 *g* for 10 min at 4°C).

   Note: This expression time is an average value, which depends on the aggregation properties of the construct and any proteolytic degradation, e.g., in linker regions of fusion proteins. Robust constructs can be expressed for longer times. Ideally, this should be checked before on a small scale.
4. Resuspend the cell pellet in 40 ml 1× MHA buffer containing 0.5 M NaCl and add Benzonase (Merck) to a final concentration of 10 U/ml for removal of nucleic acids.

   Note: To reduce protein degradation, protease inhibitors can be added to the solubilized cells. Proteolysis is usually only an issue for some fusion proteins, especially with positively charged residues in or near the linker region. It should

be kept in mind that most scFv fragments are not readily degraded by proteases. The commercial protease inhibitor cocktails are mostly targeting eukaryotic proteases and are thus not very effective against *E. coli* proteases. Also, proteolysis, if it occurs by periplasmic enzymes, frequently begins during the induction phase, and can therefore only partially be combated with inhibitors.

5. Disrupt the cells using a French Press (20,000 psi, 4°C in a cold room) or the TS 1.1 benchtop. For the French Press, perform at least three passages for optimal lysis of the cells.

   Note: The large-scale protocol consists of a lysis of the whole cells, not a periplasmic extraction. The latter can be done as an alternative, but is usually more difficult to do reproducibly on large scales.
6. Centrifuge the crude extract in order to separate insoluble cell debris from soluble protein (20,000 *g*, 30 min at 4°C). Carefully separate supernatant from pellet and transfer it to a new tube.
7. Filter the supernatant through a 0.22-µm filter (use filters with low protein binding properties, e.g., Durapore filters from Millipore). Save an aliquot for subsequent analysis by SDS-PAGE.

27.3.4 Purification of scFv Fragments

The purification scheme described below includes immobilized metal affinity chromatography (IMAC) as the main step in combination with a directly coupled ion-exchange (IEX) chromatography for separation of the scFvs from bacterial proteins. It is, in general, difficult to get a very highly pure product after a single step of IMAC. Also, such preparations frequently contain a significant amount of RNA or DNA. This motivated the use of the coupled system (Fig. 27.3)

For IEX chromatography, calculate the isoelectric point (pI) of the scFv on the basis of its amino acid composition (e.g., using the website www.expasy.org/cgi-bin/protparam), as this value is important for deciding which ion exchange matrix and buffer system to use: for scFvs with pI values below 7.0 we recommend using an anion exchanger, while for values higher than 7.0, cation exchange chromatography should be performed. Purification on the BioCAD 700E (PerSeptive Biosystems, acquired by Applied Biosystems) over both columns can be done within only 30 min either by manual operation or by running a program automatically.

Note: If possible, perform all chromatography steps at 4°C. Use only buffers of highest purity, properly degassed and filtered (0.22 µm) prior to use. The system should be completely purged with 1× MHA buffer before the start of purification to avoid any air bubbles in the tubings, which might subsequently get trapped on the columns.

1. Prepare the Ni-IDA POROS MC column (having a column volume (CV) of 1.7 ml) by preloading it with 10 CV 200 mM $NiCl_2$ and subsequent washing with 10 CV sterile distilled water to remove the excess Ni^{2+} ions. Equilibrate the column with 10 CV 1× MHA buffer, 150 mM NaCl, pH 7.0.

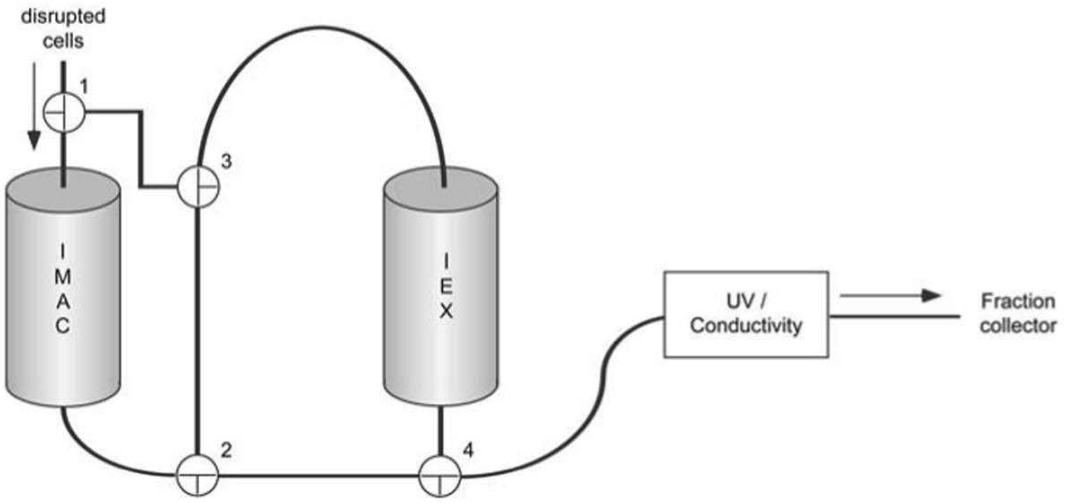


Fig. 27.3 *Tubing diagram for rapid two-column purification of antibody fragments*. The disrupted and filtrated cells are loaded onto the immobilized metal ion affinity chromatography (IMAC) column first. Upon antibody elution by increased imidazole concentration, the eluant flow is redirected onto the ion-exchange (IEX) column by turning valves 2 and 4. The adsorbed protein is finally eluted by applying a salt gradient and reversing/switching valves 1 and 3. Please note that it is essential that the imidazole used for elution does not have a high ionic strength, requiring that its pH is adjusted with acetic acid (see note in Sect. 27.2 "Materials")

Note: Similar chromatographic materials can be used with other chromatography systems.

2. Load the filtrated antibody sample onto the POROS MC column. During sample loading, the flow rate – otherwise being 3 ml/min – should be reduced to 1.5 ml per minute.
3. Wash the column with 15 CV 1 × MHA buffer containing 150 mM NaCl. The UV absorption signal at 280 nm should have reached its baseline by then.
4. Wash the column with 10 CV 1 × MHA buffer containing 30 mM NaCl, pH 7.0.
5. Wash the column with 10 CV 1 × MHA buffer containing 1 M NaCl, pH 7.0.

   Note: Washing with low and high salt concentrations assists in removing unspecifically bound material. If the protein of interest is present only in a small amount, several contaminating bacterial proteins can bind to the IMAC column under purification conditions and would finally coelute with the scFv if these stringent washing steps were omitted.
6. Wash column with 10 CV of 30 mM imidazole, 150 mM NaCl, pH 7.0.
7. Elute specifically bound scFv by either applying an imidazole gradient from 30 to 250 mM imidazole (pH 7.0) (no salt) (10 CV) or a step elution with 250 mM imidazole (pH 7.0) (no salt) (10 CV).
8. Directly load the elution on the downstream IEX column by using the BioCAD workstation or equivalent (for tubing diagram see Fig. 27.3). This column can either be a cation or an anion exchanger (see note at the beginning of this subsection).

Note: The pH for the following washing step and the final elution depends on the pI of the antibody fragment and on the type of the column used (i.e., if the antibody has a pI of 8.5, the pH should be adjusted to 7.0 and the sample should be applied to a cation exchange column; however, if the scFv fragment's pI is lower than 7.0, work with an anion exchanger at pH 8.0).

9. Wash the column with 1× MHA buffer, containing 30 mM NaCl, at the appropriate pH until the UV 280 nm baseline is reached.
10. Elute the scFv from the ion exchange column with a salt gradient from 30 to 750 mM NaCl with the appropriate pH (15 CV). Monitor the elution by its UV absorbance at 280 nm and collect 0.5 ml fractions. Analyze each of them by SDS-PAGE and pool those containing pure scFvs.

    Note: The imidazole stock solution used to elute the protein from IMAC must be pH-adjusted by using acetic acid and not with HCl, in order to keep the ionic strength low (otherwise, the protein might run through the coupled downstream ion exchange column).
11. Finally, determine the concentration of this protein solution using standard procedures, and store the purified scFv at 4°C after addition of 0.05% sodium azide. For long-term storage at −80°C, it might be beneficial to stabilize the purified scFv by adding human serum albumin to a final concentration of 10 mg/ml.

27.4 Comments

This part of the protocol contains general comments about the recommended standard method. The most critical steps were already highlighted directly following the instructions in the different subsections.

(a) The methods outlined in this chapter will almost certainly be used on antibody fragments that are intrinsically aggregation-prone, this being their main motivation. It must be kept in mind that a tendency for aggregation is an intrinsic property of the protein, and cannot be overcome upon successful expression by whatever method. Such antibody fragments tend to form *soluble* aggregates, and thus a mere inspection of soluble protein on western blots after expression may be very misleading. Molecular chaperones can prevent the formation of large, insoluble aggregates, but sometimes not of smaller, soluble aggregates. Therefore, a serious characterization of an scFv fragment must include gel chromatography, ideally coupled with multi-angle light scattering. This will give a very clear description of the amount of soluble aggregates in a preparation, or their development over time.

(b) Co-expressing Skp together with an antibody fragment might sometimes result in a prolonged lag phase and slower doubling time of the bacterial cells. However, upon reaching an OD_{600} of 0.8, these cells recover, possess a higher doubling rate, and finally lead to higher yield of recombinant protein.

(c) In contrast to the production of poorly folding antibodies in the absence of chaperone, scFv expression in their presence also offers the advantage of the ability to increase the time of expression. As chaperone expression results in less cell lysis, the final cell density can be increased, which also results in increased total scFv yield.
(d) We previously demonstrated that the co-expression of the periplasmic PPIase SurA produced no increase in the functional scFv fragment level in the periplasm, at least for the scFv fragments tested (Bothmann and Plückthun 2000; Ramm and Plückthun 2000). However, we decided to retain its gene in the pCH vector series, as we did not observe any disadvantage of SurA expression. In addition, we wanted these vectors to be as generally applicable as possible, also being able to assist the folding of proteins other than scFv.
(e) As most of the *E. coli* host proteins co-purified in IMAC have a pI of less than 6.5, they will bind to anion-exchange columns. Therefore, these columns can also be used in an inverse setup for scFv constructs with high pI, trapping the *E. coli* proteins while leaving the scFv in the flow-though.
(f) As imidazole slowly catalyzes the hydrolysis of acid labile bonds and can interfere with many subsequent assays, its presence is not ideal for long-term storage. Therefore, the two-step method presented in this protocol helps as built-in buffer exchange. Alternatively, the IMAC eluate can be dialyzed against a physiological buffer such as PBS immediately after purification.

Acknowledgments This protocol has evolved over the years, and heavily relies on the research and the original versions developed by Hendrick Bothmann.

References

Bass S, Gu Q, Christen A (1996) Multicopy suppressors of prc mutant *Escherichia coli* include two HtrA (DegP) protease homologs (HhoAB), DksA, and a truncated R1pA. J Bacteriol 178:1154–1161

Bothmann H, Plückthun A (1998) Selection for a periplasmic factor improving phage display and functional periplasmic expression. Nat Biotechnol 16:376–380

Bothmann H, Plückthun A (2000) The periplasmic *Escherichia coli* peptidylprolyl cis, trans-isomerase FkpA I. Increased functional expression of antibody fragments with and without cis-prolines. J Biol Chem 275:17100–17105

Chen J, Song JL, Zhang S, Wang Y, Cui DF, Wang CC (1999) Chaperone activity of DsbC. J Biol Chem 274:19601–19605

De Cock H, Schäfer U, Potgeter M, Demel R, Müller M, Tommassen J (1999) Affinity of the periplasmic chaperone Skp of *Escherichia coli* for phospholipids, lipopolysaccharides and non-native outer membrane proteins. Role of Skp in the biogenesis of outer membrane protein. Eur J Biochem 259:96–103

Ewert S, Honegger A, Plückthun A (2003) Structure-based improvement of the biophysical properties of immunoglobulin VH domains with a generalizable approach. Biochemistry 42:1517–1528

Ewert S, Honegger A, Plückthun A (2004) Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering. Methods 34:184–199

Galat A (2003) Peptidylprolyl cis/trans isomerases (immunophilins): biological diversity–targets–functions. Curr Top Med Chem 3:1315–1347

Honegger A, Malebranche AD, Röthlisberger D, Plückthun A (2009) The influence of the framework core residues on the biophysical properties of immunoglobulin heavy chain variable domains. Protein Eng Des Sel 22:121–134

Hu X, O'Hara L, White S, Magner E, Kane M, Wall JG (2007) Optimisation of production of a domoic acid-binding scFv antibody fragment in *Escherichia coli* using molecular chaperones and functional immobilisation on a mesoporous silicate support. Protein Expr Purif 52:194–201

Jarchow S, Lück C, Görg A, Skerra A (2008) Identification of potential substrate proteins for the periplasmic *Escherichia coli* chaperone Skp. Proteomics 8:4987–4994.

Jermutus L, Honegger A, Schwesinger F, Hanes J, Plückthun A (2001) Tailoring *in vitro* evolution for protein affinity or stability. Proc Natl Acad Sci USA 98:75–80

Jung S, Plückthun A (1997) Improving in vivo folding and stability of a single-chain Fv antibody fragment by loop grafting. Protein Eng 10:959–966

Jung S, Honegger A, Plückthun A (1999) Selection for improved protein stability by phage display. J Mol Biol 294:163–180

Kadokura H, Katzen F, Beckwith J (2003) Protein disulfide bond formation in prokaryotes. Annu Rev Biochem 72:111–135

Knappik A, Plückthun A (1995) Engineered turns of a recombinant antibody improve its in vivo folding. Protein Eng 8:81–89

Kolaj O, Spada S, Robin S, Wall JG (2009) Use of folding modulators to improve heterologous protein production in *Escherichia coli*. Microbial Cell Factories 8:9

Krebber A, Bornhauser S, Burmester J, Honegger A, Willuda J, Bosshard HR, Plückthun A (1997) Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. J Immunol Methods 201:35–55

Kügler M, Stein C, Schwenkert M, Saul D, Vockentanz L, Huber T, Wetzel SK, Scholz O, Plückthun A, Honegger A, Fey GH (2009) Stabilization and humanization of a single-chain Fv antibody fragment specific for human lymphocyte antigen CD19 by designed point mutations and CDR-grafting onto a human framework. Protein Eng Des Sel 22:135–147

Levy R, Weiss R, Chen G, Iverson BL, Georgiou G (2001) Production of correctly folded Fab antibody fragment in the cytoplasm of *Escherichia coli* trxB gor mutants via the co-expression of molecular chaperones. Protein Expr Purif 23:338–347

Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1–I2 regulatory elements. Nucleic Acids Res 25:1203–1210

Maurer R, Meyer B, Ptashne M (1980) Gene regulation at the right operator (OR) bacteriophage lambda I. OR3 and autogenous negative control by repressor. J Mol Biol 139:147–161

Nieba L, Honegger A, Krebber C, Plückthun A (1997) Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. Protein Eng 10:435–444

Ortenberg R, Beckwith J (2003) Functions of thiol-disulfide oxidoreductases in *E. coli*: redox myths, realities, and practicalities. Antioxid Redox Signal 5:403–411

Plückthun A, Krebber A, Krebber C, Horn U, Knüpfer U, Wenderoth R, Nieba L, Proba K, Riesenberg D (1996) Producing antibodies in *Eschericia coli*: Fom PCR to fermentation. In: McCafferty J, Hoogenboom H (eds) Antibody engineering: a practical approach. IRL press, Oxford, pp 203–252

Proba K, Ge L, Plückthun A (1995) Functional antibody single-chain fragments from the cytoplasm of *Escherichia coli*: Influence of thioredoxin reductase (TrxB). Gene 159:203–207

Proba K, Wörn A, Honegger A, Plückthun A (1998) Antibody scFv fragments without disulfide bonds made by molecular evolution. J Mol Biol 275:245–253

Ramm K, Plückthun A (2000) The periplasmic *Escherichia coli* peptidylprolyl cis, trans-isomerase FkpA II. Isomerase-independent chaperone activity in vitro. J Biol Chem 275: 17106–17113

Sambrook J, Russell D (2001) Molecular cloning. A laboratory manual, 3rd edn. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press 2001

Sandee D, Tungpradabkul S, Kurokawa Y, Fukui K, Takagi M (2005) Combination of Dsb co-expression and an addition of sorbitol markedly enhanced soluble expression of single-chain Fv in *Escherichia coli*. Biotechnol Bioeng 91:418–424

Schäfer U, Beck K, Müller M (1999) Skp, a molecular chaperone of gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins. J Biol Chem 274:24567–24574

Schimmele B, Plückthun A (2005) Engineering proteins for stability and efficient folding. In: Buchner J, Kiefhaber T (eds) Protein folding handbook. Wiley Verlag GmbH & Co. KGaA, Weinheim, Germany, pp 1281–1333

Schlapschy M, Grimm S, Skerra A (2006) A system for concomitant overexpression of four periplasmic folding catalysts to improve secretory protein production in *Escherichia coli*. Protein Eng Des Sel 19:385–390

Segatori L, Paukstelis PJ, Gilbert HF, Georgiou G (2004) Engineered DsbC chimeras catalyze both protein oxidation and disulfide-bond isomerization in *Escherichia coli*: reconciling two competing pathways. Proc Natl Acad Sci USA 101:10018–10023

Skerra A, Plückthun A (1988) Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli*. Science 240:1038–1041

Skorko-Glonek J, Sobiecka-Szkatula A, Narkiewicz J, Lipinska B (2008) The proteolytic activity of the HtrA (DegP) protein from *Escherichia coli* at low temperatures. Microbiology 154:3649–3658

Söderlind E, Dueñas M, Borrebaeck CA (1995) Chaperonins in phage display of antibody fragments. Methods Mol Biol 51:343–353

Thorpe SJ, Kerr MA (1994) Common immunological techniques: ELISA, blotting, immunohistochemistry and immunocytochemistry. In: Kerr M, Thorpe R (eds) Immunochemistry. Oxford, Labfax, BIOS Scientific Publishers Limited, pp 175–209

Velappan N, Sblattero D, Chasteen L, Pavlik P, Bradbury AR (2007) Plasmid incompatibility: more compatible than previously thought. Protein Eng Des Sel 20:309–313

Wall JG, Plückthun A (1995) Effects of overexpressing folding modulators on the *in vivo* folding of heterologous proteins in *Escherichia coli*. Curr Opin Biotechnol 6:507–516

Willuda J, Honegger A, Waibel R, Schubiger PA, Stahel R, Zangemeister-Wittke U, Plückthun A (1999) High thermal stability is essential for tumor targeting of antibody fragments: engineering of a humanized anti-epithelial glycoprotein-2 (epithelial cell adhesion molecule) single-chain Fv fragment. Cancer Res 59:5758–5767

Wörn A, Plückthun A (2001) Stability engineering of antibody single-chain F_v fragments. J Mol Biol 305:989–1010

Wörn A, Aufder Maur A, Escher D, Honegger A, Barberis A, Plückthun A (2000) Correlation between in vitro stability and in vivo performance of anti-GCN4 intrabodies as cytoplasmic inhibitors. J Biol Chem 275:2795–2803

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119

Zhao Z, Peng Y, Hao S-F, Zeng Z-H, Wang C-C (2003) Dimerization by domain hybridization bestows chaperone and isomerase activities. J Biol Chem 278:43292–43298

4. General Discussion and Outlook

The main objective of this thesis was to investigate engineering approaches for improving the biophysical properties of full-length IgG antibodies, mainly focusing on stability and aggregation-related aspects. Similar analyses have already been performed in the past - however, those studies were mainly restricted to smaller antibody fragments like scFv or Fab produced in prokaryotic systems due to various reasons. The first and foremost was that the initial analyses had to be able to untangle effects in the domain itself and on interactions with other domains. In addition, it had to be ensured that the effects could be observed at all - a task that becomes progressively more difficult with more domains. Thus, this question would have been unanswerable with a full-length IgG initially. Now, having the data for single domains and scFv fragments at our disposal, the question can be posed whether similar effects are seen in IgGs - the antibody format mainly used for therapeutical antibodies nowadays.

It could be shown that enhanced stabilities both with respect to thermal and denaturant-induced unfolding can be achieved by rational engineering for various antibody formats and independent of the expression system. The presented data thus clearly indicate the transferability of improvements implemented in smaller antibody fragments onto full-lengths IgGs - a finding having important consequences: since the variable domains of IgGs and their features can be modified in smaller fragments, the required experimental set-ups can be simplified with respect to the employed format and the use of a prokaryotic system. Importantly, the findings presented here not only allow to retrospectively improve existing binders, but above all might also have important implications for future library designs. So far, based on the stability of the individual domains determined in isolation and within scFv fragments, it could be assumed that high-quality libraries for the selection of antibodies should only contain molecules possessing the most stable V_H and V_L domains, namely V_H3 and $V_κ3$, respectively. The underlying rationale is that most applications prefer or even require the resulting antibodies to be stable and to be obtained as soluble proteins in high yields. However, these libraries would have the great disadvantage of being limited to only one framework combination, thereby reducing the binding diversity of its members significantly. As known from several analyses, the framework residues located either at the hapten binding interface or the outer loop greatly contribute to both affinity and diversity, but are rather variable among the different V_H families. Therefore, it is not advisable to limit the members of the initial library to a few or even one framework only. So far, selected binders possessing a framework combination different from the best one might have suffered from suboptimal biophysical properties, like reduced stability or enhanced aggregation susceptibility. Consequently, it would be best for future library designs to create structurally diverse libraries of stable frameworks after optimizing the used consensus antibody frameworks in the first place. Since the study presented here restricted itself to members of the least stable V_H6 family, it might be of interest to perform similar engineering experiments with members of the other even-numbered V_H domains known for their less favorable biophysical properties. Although based on the sequence similarities presumably analogous results are to be expected for domains belonging to the V_H2 and V_H4 family, it

would be beneficial to have advanced consensus sequences available for these families as well. The study presented in this thesis should be of great help at this as it contributes to the understanding of the design of such consensus sequences and highlights promising procedures to improve them.

Interestingly, the described mutations not only influenced the biophysical properties of the engineered full-length IgG but also had an impact on their structural integrity and homogeneity. It would be of great significance if IgGs with increased homogeneity had benefits over their heterologous counterparts, especially in the context of clinical applications with their potential side-effects. If few amino acid exchanges could enhance the efficiency and safety of antibody drugs this would not only present crucial benefits for the pharmaceutical industry (having better products that potentially require less laborious downstream processing) but, more importantly, for the patients as well. Of particular interest is furthermore the presented finding that the increased homogeneity was not only restricted to the V_H domain but even affected the structural integrity of the adjacent C_H1 domain with its intrachain disulfide bond. It is generally assumed that each of the IgG domains acts as an independent unit, especially being autonomous regarding its folding. However, it has already been demonstrated that adjacent domains can influence their neighbors by longitudinal interactions.[1] While this intramolecular signaling has been the subject of controversial debates, the presented data clearly indicate that one domain might not only influence the stability but also the structure of another domain. To finally reveal the underlying mode of interaction, molecular details of the domain structures would be of enormous advantage. Therefore, crystal structures of both the WT and M variants should be prepared and analyzed with respect to differences in the V_H-V_L-interface or in the domain structures itself. Since antibodies have been successfully studied by X-ray crystallography for several decades,[2; 3] this should be feasible and be performed – especially, as the conclusions of such investigations might be of great value for the future design of antibodies and for their various applications.

Another main aspect of this thesis was the comparison of different eukaryotic expression systems, namely HEK293 cells and *Pichia pastoris*. In recent years, this yeast system has enjoyed an increasing popularity and has proven to be a very useful platform for the production of various proteins. The outlined results clearly indicate that *Pichia* is also well suited for the expression of full-lengths IgG molecules at a large scale. Using the constitutively active GAP promoter and the proposed sequence designs, it is possible to continuously produce genuine antibodies without any remaining residues from the sequences used for secretion of the molecules. This should - combined with glycoengineered *Pichia* strains as described below - be quite helpful to further widen the application range of this expression system. Alternatively, the production of IgGs with reduced aggregation propensities can now easily be performed by intentionally using the α-factor pre-pro sequence, resulting in EAEA-tetrapeptides still being present at the respective N-termini after enzymatic cleavage.

The usage of the *Pichia* system revealed another surprising feature. The presented results clearly indicate that the *Pichia* glycan moiety - despite slightly destabilizing the C_H2 domain - has the surprising effect of protecting IgGs against aggregation. All analyzed glycosylated *Pichia*-produced IgGs showed a higher resistance to aggregation compared to their counterparts expressed in

mammalian cells. Constructs with additional sequences remaining from the α-factor pre-pro sequence even did not exhibit any aggregation at all under the conditions tested. Since the impact of the glycans seems not to be based on their contribution to the protein's overall charge (as shown by IEF), this phenomenon seems to be independent of the protein's isoelectric point and thus probably should be transferable to other molecules as well. Therefore, it would be of great interest to elucidate which modules of the *Pichia* glycan are accountable for this characteristic feature and to also examine the glycan's impact on proteins other than IgGs. Recently, glycoengineered *Pichia* strains have been developed and meanwhile successfully applied to the production of monoclonal antibodies.[4-7] In this method based on the GlycoSwitch technology, already existing expression strains are gradually modified by the disruption of an endogenous glycosyltransferase gene followed by the stepwise introduction of heterologous glycosylation enzymes. Each of these engineering steps equips the *Pichia* endoplasmic reticulum and Golgi system with additional glycosidase or glycosyltransferase activities, resulting in individual strains expressing proteins with different intermediate glycoforms. Eventually, all steps combined cause the engineered strain to produce human-like $(Gal)_2(GlcNAc)_2(Man)_3(GlcNAc)_2$ N-glycans. To shed light on the question which components of the original *Pichia* glycan are responsible for its aggregation-preventing properties, IgG antibodies possessing intermediate glycoforms should be produced and subsequently analyzed for their aggregation susceptibilities. The finding might even lead to a better understanding of the aggregation mechanisms of antibodies in general and may present further engineering methods to successfully tackle this problem. The resulting insight could potentially even facilitate additional applications of aggregation-resistant antibodies in the future.

As already mentioned, also the EAEA-tetrapeptides left behind from the α-factor pre-pro sequence had a dramatic positive impact on the aggregation propensities of *Pichia*-produced IgGs, indicating that this susceptibility can be engineered by modifying the charge distributions at certain locations. Although the presence of these amino-terminal extensions might not always be desirable or could even potentially create new immunogenic epitopes in human therapy, these appendices may still be of great advantage in various other applications. For example, these four amino acids should not exhibit any drawback in *in vitro* usage of modified antibodies in the respective set-ups - like the detection of different proteins on ELISA or immunoblotting - since unaltered affinities were shown in this work. As a major advantage, the strategy of attaching the charged EAEA-peptides to the N-terminus of HEK-produced IgGs is easily applicable to a large variety of molecules, because it can readily be performed by simple cloning means. Thus, it could also be applied retrospectively to already existing binders without any difficulty. Importantly, this approach may supersede the elsewhere outlined idea of introducing charged residues into the CDR regions.[8] Since that method may abolish the antibody-antigen recognition directly or indirectly by altering the loop conformation, it cannot be considered a generally applicable strategy. In further studies it should be investigated whether the introduction of additional charges at other locations within the IgG molecule (like, for example, at the C-terminus or in the hinge region of the molecule) would have a similar effect as the N-terminal EAEA-peptide. In these studies, care should be taken to also analyze other properties of the resulting molecules like thermal stability and expression yields.

Furthermore, it is rather unlikely that the described EAEA-tetrapeptide already constitutes the ultimate appendix to antibody molecules with respect to reduced aggregation propensities. As it could be shown that the addition of various peptides to the N-termini only marginally influenced the expression level of IgGs expressed in mammalian cells, peptides other than EAEA should be analyzed for their feasibility and influence. Therefore, obtaining more insight into the exact mode of action as well as detailed knowledge derived from crystal structures would be of great benefit. Alternatively, small peptide libraries may be attached to the antibody and screened for their impact onto the aggregation susceptibility if appropriate selection methods can be developed and implemented.

As summarized so far, this study analyzed the impact of various amino acid alterations on the biophysical characteristics of different antibodies. The results might help to enhance the properties of already existing antibodies and to initially improve library members, allowing the selection of more stable binders or molecules that can act under harsh conditions such as the presence of denaturants. It would further be of great interest to use the knowledge obtained and the broad range of available IgG variants - stability engineered mutants, truncated variants as well as glycan-knockouts - to gain additional insight into kinetic aspects of their stability and to determine how this feature is influenced by various factors. In most studies, standard analyses of thermal stabilities are oversimplified as they are performed by equilibrium thermodynamics. While this strategy might be justified in some cases, it may not be the most adequate procedure for antibodies due to their multidomain structure and the irreversible nature of their thermally-induced denaturation. Therefore, it would be more accurate to use a comprehensive approach and to conduct a set of thermal-denaturation experiments of constructs with various combinations of domains at different scan rates and varying protein concentrations. As a drawback, such an extensive study requires large amounts of purified proteins that themselves demand appropriate mammalian expression facilities. The method of choice for the necessary biophysical analyses is Differential Scanning Calorimetry (DSC) since it not only enables the determination of thermodynamic features but also indirectly of kinetic parameters that are necessary for characterizing the overall stability of antibodies. As shown in this thesis, an important advantage of DSC is its ability to distinguish different transitions in multidomain proteins such as IgGs that are often overlapping in e.g. spectroscopic methods. Recent improvements in high-throughput DSC and in mathematical models facilitate the determination of the role of factors on protein stability such as concentration and storage conditions. It would be a major benefit to find correlations (if they exist at all) between thermodynamic and kinetic stabilities of individual domains and within the whole protein. Eventually, kinetic analysis of thermal denaturation of IgGs will provide proper parameters for the determination of important characteristics like their half-life (shelf-life) as a function of storage temperature. In addition, the influence of various additives or osmolytes on the stabilization of IgGs could be investigated by this set-up. These small molecules, ranging from sugars over polyols to amino acids and their derivatives, have been shown to counteract various stress conditions of proteins[9; 10] and thus are used for the development of optimal storage conditions.[11] Incorporating these aspects in our analyses, it should be possible to get a deeper insight into the exact mechanism by which this stabilization occurs and to compare the effects of various osmolytes. The resulting

conclusions will be of great importance for improving our understanding of antibody stability and even may pinpoint new hotspots of future engineering approaches.

Taken together, the analyses presented in this thesis have dealt with various aspects of antibodies and with some of the challenges that still derogate the successful widening of their application range. With the help of the outlined modifications and the acquired insight, antibody engineering should be significantly pushed forward and may result in an advanced and improved IgG antibody generation, thereby perpetuating the success story of these molecules in the 21st century.

References

1. Röthlisberger, D., Honegger, A. & Plückthun, A. (2005). Domain interactions in the Fab fragment: a comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability. *J Mol Biol* **347**, 773-789.
2. Silverton, E. W., Navia, M. A. & Davies, D. R. (1977). Three-dimensional structure of an intact human immunoglobulin. *Proc Natl Acad Sci U S A* **74**, 5140-5144.
3. Harris, L. J., Skaletsky, E. & McPherson, A. (1998). Crystallographic structure of an intact IgG1 monoclonal antibody. *J Mol Biol* **275**, 861-872.
4. Li, H., Sethuraman, N., Stadheim, T. A., Zha, D., Prinz, B., Ballew, N., Bobrowicz, P., Choi, B. K., Cook, W. J., Cukan, M., Houston-Cummings, N. R., Davidson, R., Gong, B., Hamilton, S. R., Hoopes, J. P., Jiang, Y., Kim, N., Mansfield, R., Nett, J. H., Rios, S., Strawbridge, R., Wildt, S. & Gerngross, T. U. (2006). Optimization of humanized IgGs in glycoengineered Pichia pastoris. *Nat. Biotechnol.* **24**, 210-215.
5. Jacobs, P. P., Geysens, S., Vervecken, W., Contreras, R. & Callewaert, N. (2009). Engineering complex-type N-glycosylation in Pichia pastoris using GlycoSwitch technology. *Nat. Protoc.* **4**, 58-70.
6. Potgieter, T. I., Cukan, M., Drummond, J. E., Houston-Cummings, N. R., Jiang, Y., Li, F., Lynaugh, H., Mallem, M., McKelvey, T. W., Mitchell, T., Nylen, A., Rittenhour, A., Stadheim, T. A., Zha, D. & d'Anjou, M. (2009). Production of monoclonal antibodies by glycoengineered Pichia pastoris. *J. Biotechnol.* **139**, 318-325.
7. Vervecken, W., Callewaert, N., Kaigorodov, V., Geysens, S. & Contreras, R. (2007). Modification of the N-glycosylation pathway to produce homogeneous, human-like glycans using GlycoSwitch plasmids. *Methods Mol. Biol.* **389**, 119-138.
8. Perchiacca, J. M., Bhattacharya, M. & Tessier, P. M. (2011). Mutational analysis of domain antibodies reveals aggregation hotspots within and near the complementarity determining regions. *Proteins* **79**, 2637-2647.
9. Kumar, R. (2009). Role of naturally occurring osmolytes in protein folding and stability. *Arch Biochem Biophys* **491**, 1-6.
10. Hamada, H., Arakawa, T. & Shiraki, K. (2009). Effect of additives on protein aggregation. *Curr Pharm Biotechnol* **10**, 400-407.
11. Bhambhani, A., Kissmann, J. M., Joshi, S. B., Volkin, D. B., Kashi, R. S. & Middaugh, C. R. (2011). Formulation design and high-throughput excipient selection based on structural integrity and conformational stability of dilute and highly concentrated IgG1 monoclonal antibody solutions. *J Pharm Sci*, Epub ahead of print.

5. Appendix

5.1 Abbreviations

αMFpp	α-factor pre-pro sequence
ADCC	Antibody-dependent cellular cytotoxicity
CD	Circular dichroism
CDC	Complement-dependent cytotoxicity
CDR	Complementarity determining region
CHO	Chinese hamster ovary cells
CIEX	Cation exchange chromatography
CMV	Cytomegalovirus
DMEM	Dulbecco's modified Eagle medium
DSC	Differential scanning calorimetry
DSF	Differential scanning fluorimetry
DTT	Dithiothreitol
EC	*Escherichia coli*
ELISA	Enzyme-linked immunosorbent assay
Fab	Fragment antigen binding
Fc	Fragment crystallizable
FcγR	Fc gamma receptors
FBS	Fetal bovine serum
GdnHCl	Guanidine hydrochloride
H	Heavy chain of an antibody
HEK	Human embryonic kidney cells
HPLC	High-performance liquid chromatography
HuCAL	Human Combinatorial Antibody Library
HT	High tension voltage
IEF	Isoelectric focusing
IPTG	Isopropyl-β-D-thiogalactopyranoside
IgG	Immunoglobulin G
ITF	Intrinsic tryptophane fluorescence
L	Light chain of an antibody
M	Mutant
mAb	Monoclonal antibodies
MALS	Static multi-angle light scattering
MHC	Major histocompatibility complex
MRE	Mean residue ellipticity
MS	Mass spectrometry

MST	Microscale thermophoresis
PBS	Phosphate buffered saline
pI	Isoelectric point
POI	Protein of interest
PP	*Pichia pastoris*
scFv	Single-chain variable fragment of an antibody
SEC	Size-exclusion chromatography
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
V_H	Variable domain of the heavy chain of an antibody
V_L	Variable domain of the light chain of an antibody
WT	Wild type

5.2 Plasmid / Vector overviews

For the expression of full-length IgG or Fab molecules, different vectors were used depending on the expression system. Here, representative vectors are shown for the IgG 2C2 WT and Fab 2C2 WT variants. Vectors encoding the M variants or the 6B3 constructs were constructed in an analogous manner. In general, the variable V_H and V_L domains, differing between the various antibodies, are shown in blue. The corresponding promoters used for the expression of the heavy and light chain are depicted in green while the origins of replication utilized for plasmid maintenance in *E. coli* are presented in red. As the Fab fragments cannot be purified from the supernatant via Protein A affinity chromatography, their Fd fragment, consisting of V_H and C_H1, possesses both an additional myc-tag (EQKLISEEDL) and a hexa his-tag at its C-terminus.

Mammalian vectors:

For the establishment of stable HEK293 cell lines expressing the various antibody formats, derivatives of the pcDNA5 vector (Invitrogen, Fig. 5.1) were used. Two complete expression cassettes for the heavy and light chain, respectively, were combined in a single vector containing constitutively active CMV promoters upstream of the endogenous IgG signal sequences.

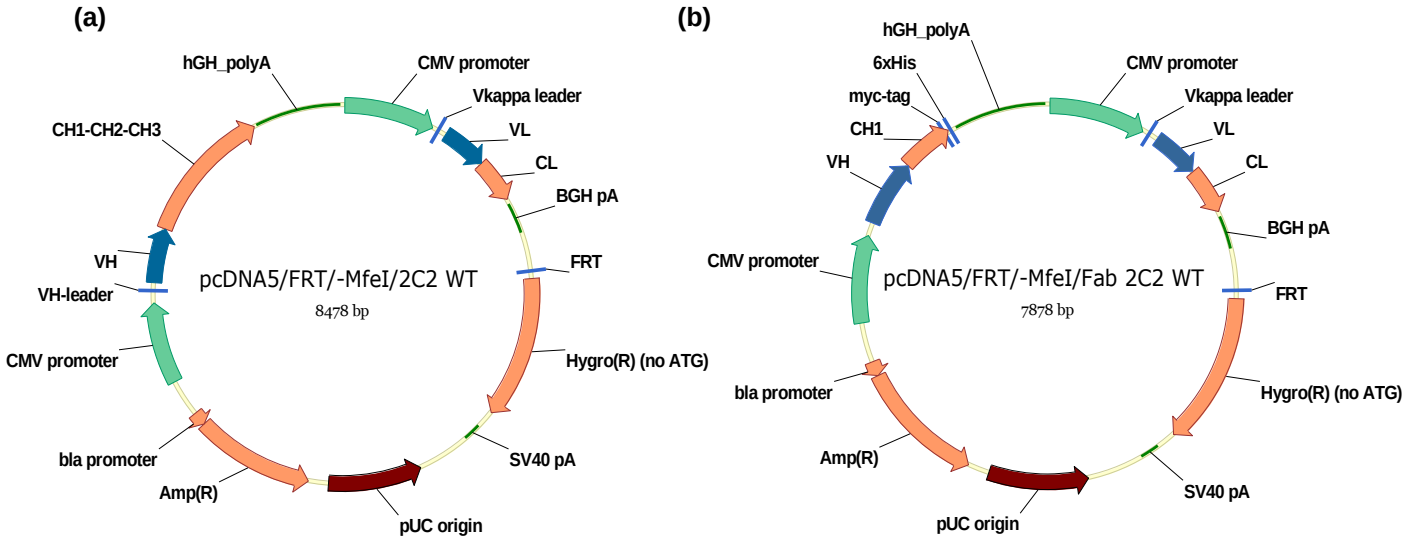


Fig. 5.1: Vector maps of pcDNA5/FRT-derivatives for the expression of (a) IgG and (b) Fab fragments in stable HEK293 cells. *Vkappa leader, VL, CL*: sequence coding for variable and constant domain of the light chain; *VH leader, VH, CH1, CH1-CH2-CH3*: sequence coding for variable and constant domain(s) of the heavy chain; *myc-tag*: sequence coding for the c-myc tag; *6xHis*: sequence coding for the hexa his-tag; *CMV promotor*: constitutively active promoter for the expression of heavy and light chain; *Hygro(R)*: gene coding for hygromycin B resistance (hygromycin B phosphotransferase); *Amp(R):* gene coding for ampicillin resistance (β-lactamase); *bla promoter:* β-lactamase promoter; *pUC origin*: origin of replication for maintenance in *E. coli*; *FRT*: FLP recombinase target, used for integration into the HEK genome; *BGH pA*: poly-adenylation site (bovine growth hormone poly-A), used for the light chain; *hGH_polyA*: poly-adenylation site (human growth hormone poly-A), used for heavy chain construct; SV40 pA: polyadenylation site (simian virus 40 poly-A), used for hygromycin B phosphotransferase.

Pichia vectors:

Expression of IgG or Fab fragments in *Pichia pastoris* was achieved after integration of derivatives of the pGAPZαB vectors (Fig. 5.2) into the yeast genome. As for HEK expression, the genes for the heavy and light chain were combined within one single vector, containing these genes as independent expression cassettes with their own promoter, their own alpha factor pre-pro sequence as well as their own terminator sequences.

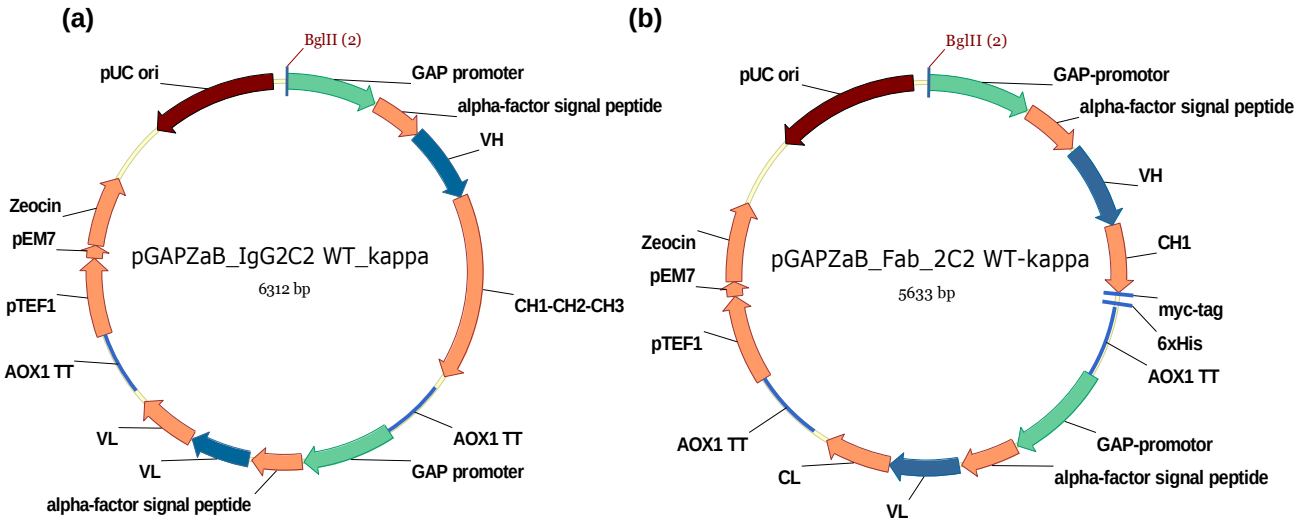


Fig. 5.2: Vector maps of pGAPZ*α*B derivatives for the (a) IgG and (b) Fab expression in *Pichia pastoris*. *VL, CL*: variable and constant domain of the light chain; *VH, CH1, CH1-CH2-CH3:* variable and constant domain(s) of the heavy chain; *alpha-factor signal peptide*: pre-pro sequence of the alpha factor for secretion; *myc-tag*: sequence coding for the c-myc tag; *6xHis*: sequence coding for the hexa his-tag; *AOX1TT*: termination sequence for light and heavy chain, respectively; *GAP-promoter*: GAP promoter; *pTEF1*: *S. cerevisiae* promoter for Zeocin™ selection in yeast; *pEM7* : bacterial promoter for Zeocin™ selection in *E. coli*; *pUC ori*: origin of replication for maintenance in *E. coli*; *Zeocin*: Zeocin resistance gene; *BglII: Bgl*II restriction site for linearization of the final vector.

Vectors for prokaryotic expression:

Prokaryotic assembly of IgG or Fab fragments was performed in the periplasm of *E. coli* SB536. Therefore, the IgG or Fab genes were arranged as a bicistronic unit, equipping each chain with its own signal sequence. The vectors used for the expression of the antibody constructs in this thesis were based on either pMx7 (for the IgG expression) or pMx9 (for the Fab production), respectively (MorphoSys, Fig. 5.3), differing in the signal sequences used for the corresponding light chain.

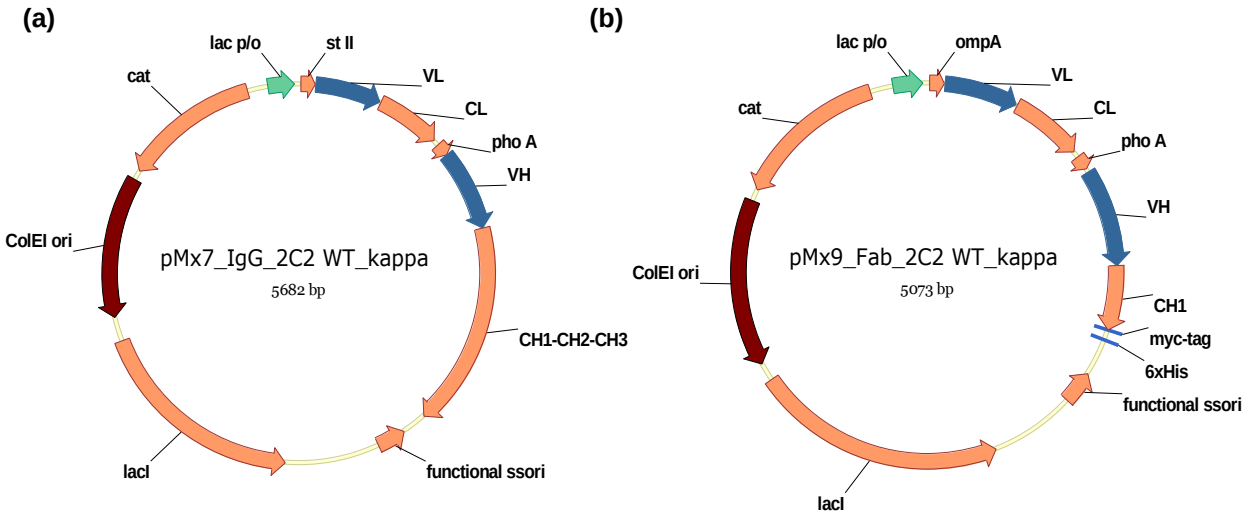


Fig. 5.3: Vector maps of pMx7 or pMx9 used for prokaryotic antibody expression. *VL, CL*: gene coding for the light chain; *VH, CH1, CH1-CH2-CH3*: gene coding for the heavy chain; *myc-tag*: sequence coding for the c-myc tag; *6xHis*: sequence coding for the hexa his-tag; *functional ssori*: functional ssDNA-Polymerase origin; *ompA*: OmpA signal sequence for periplasmic transport of the light chain; *stII*: STII signal sequence for periplasmic transport of the light chain; *phoA*: PhoA signal sequence for periplasmic transport of the heavy chain; *lac p/o*: IPTG-inducible lac promoter/operator; *cat*: chloramphenicol acetyltransferase (chloramphenicol resistance gene) ; *lacI* : lac repressor gene, active in the absence of IPTG; *ColE1 ori*: origin of replication.

5.3 Protein Sequences

The following sequences are derived from the vectors shown above and represent the translated, but yet unprocessed, proteins. Thus, signal sequences and C-terminal Lys residues subsequently cleaved off are indicated. Apart from the prokaryotic constructs - that are expressed in a bicistronic unit - the individual heavy and light chain sequences are displayed, respectively. In general, the six distinct point mutations (Q5V, S16G, T58I, V72D, S76G and S90Y) discriminating WT and M constructs are highlighted either in red for the WT or in blue for the M. Secretion signals are each identified by gray highlights while spacer peptides used in the *Pichia* system are shown in light gray. Truncated C-terminal lysines are placed into brackets. For the Fab fragments, the additional myc- and hexa-his-tags are emphasized by dark green highlights. The residue 299 that is mutated to alanine in the glycan knock-out T299A variants is italicized and underlined in the HEK and PP IgG constructs.

Mammalian (HEK) constructs:

HEK IgG 2C2 WT

MKHLWFFLLL VAAPRWVLSQ VQLQQSGPGL VKPSQTLSLT CAISGDSVSS NSAAWNWIRQ SPGRGLEWLG
RTYYRSKWYN DYAVSVKSRI TINPDTSKNQ FSLQLNSVTP EDTAVYYCAR QRGHYGKGYK GFNSGFFDFW
GQGTLVTVSS ASTKGPSVFP LAPSSKSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAVLQSS
GLYSLSSVVT VPSSSLGTQT YICNVNHKPS NTKVDKRVEP KSCDKTHTCP PCPAPELLGG PSVFLFPPKP
KDTLMISRTP EVTCVVVDVS HEDPEVKFNW YVDGVEVHNA KTKPREEQYN S*T*YRVVSVLT VLHQDWLNGK
EYKCKVSNKA LPAPIEKTIS KAKGQPREPQ VYTLPPSREE MTKNQVSLTC LVKGFYPSDI AVEWESNGQP
ENNYKTTPPV LDSDGSFFLY SKLTVDKSRW QQGNVFSCSV MHEALHNHYT QKSLSLSPG(K)

HEK IgG 2C2 M

MKHLWFFLLL VAAPRWVLSQ VQLVQSGPGL VKPGQTLSLT CAISGDSVSS NSAAWNWIRQ SPGRGLEWLG
RIYYRSKWYN DYADSVKGRI TINPDTSKNQ FYLQLNSVTP EDTAVYYCAR QRGHYGKGYK GFNSGFFDFW
GQGTLVTVSS ASTKGPSVFP LAPSSKSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAVLQSS
GLYSLSSVVT VPSSSLGTQT YICNVNHKPS NTKVDKRVEP KSCDKTHTCP PCPAPELLGG PSVFLFPPKP
KDTLMISRTP EVTCVVVDVS HEDPEVKFNW YVDGVEVHNA KTKPREEQYN S*T*YRVVSVLT VLHQDWLNGK
EYKCKVSNKA LPAPIEKTIS KAKGQPREPQ VYTLPPSREE MTKNQVSLTC LVKGFYPSDI AVEWESNGQP
ENNYKTTPPV LDSDGSFFLY SKLTVDKSRW QQGNVFSCSV MHEALHNHYT QKSLSLSPG(K)

HEK Fab 2C2 WT

MKHLWFFLLL VAAPRWVLSQ VQLQQSGPGL VKPSQTLSLT CAISGDSVSS NSAAWNWIRQ SPGRGLEWLG
RTYYRSKWYN DYAVSVKSRI TINPDTSKNQ FSLQLNSVTP EDTAVYYCAR QRGHYGKGYK GFNSGFFDFW
GQGTLVTVSS ASTKGPSVFP LAPSSKSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAVLQSS
GLYSLSSVVT VPSSSLGTQT YICNVNHKPS NTKVDKRVEP KSCDKTHLEQ KLISEEDLNS AVDHHHHHH

HEK Fab 2C2 M

MKHLWFFLLL VAAPRWVLSQ VQLVQSGPGL VKPGQTLSLT CAISGDSVSS NSAAWNWIRQ SPGRGLEWLG
RIYYRSKWYN DYADSVKGRI TINPDTSKNQ FYLQLNSVTP EDTAVYYCAR QRGHYGKGYK GFNSGFFDFW
GQGTLVTVSS ASTKGPSVFP LAPSSKSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAVLQSS
GLYSLSSVVT VPSSSLGTQT YICNVNHKPS NTKVDKRVEP KSCDKTHLEQ KLISEEDLNS AVDHHHHHH

HEK IgG 2C2 kappa

MVLQTQVFIS LLLWISGAYG DIVLTQSPAT LSLSPGERAT LSCRASQSVS SSYLAWYQQK PGQAPRLLIY
GASSRATGVP ARFSGSGSGT DFTLTISSLE PEDFATYYCQ QYYNIPTFGQ GTKVEIKRTV AAPSVFIFPP
SDEQLKSGTA SVVCLLNNFY PREAKVQWKV DNALQSGNSQ ESVTEQDSKD STYSLSSTLT LSKADYEKHK
VYACEVTHQG LSSPVTKSFN RGEC

HEK IgG 6B3 WT

MKHLWFFLLL VAAPRWVLSQ VQLQQSGPGL VKPSQTLSLT CAISGDSVSS NSAAWNWIRQ SPGRGLEWLG
RTYYRSKWYN DYAVSVKSRI TINPDTSKNQ FSLQLNSVTP EDTAVYYCAR SYFISFFSFD YWGQGTLVTV
SSASTKGPSV FPLAPSSKST SGGTAALGCL VKDYFPEPVT VSWNSGALTS GVHTFPAVLQ SSGLYSLSSV
VTVPSSSLGT QTYICNVNHK PSNTKVDKRV EPKSCDKTHT CPPCPAPELL GGPSVFLFPP KPKDTLMISR
TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ YNS*T*YRVVSV LTVLHQDWLN GKEYKCKVSN
KALPAPIEKT ISKAKGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP
PVLDSDGSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNH YTQKSLSLSP G(K)

HEK IgG 6B3 M

MKHLWFFLLL VAAPRWVLSQ VQLVQSGPGL VKPGQTLSLT CAISGDSVSS NSAAWNWIRQ SPGRGLEWLG
RIYYRSKWYN DYADSVKGRI TINPDTSKNQ FYLQLNSVTP EDTAVYYCAR SYFISFFSFD YWGQGTLVTV
SSASTKGPSV FPLAPSSKST SGGTAALGCL VKDYFPEPVT VSWNSGALTS GVHTFPAVLQ SSGLYSLSSV
VTVPSSSLGT QTYICNVNHK PSNTKVDKRV EPKSCDKTHT CPPCPAPELL GGPSVFLFPP KPKDTLMISR
TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ YNS*T*YRVVSV LTVLHQDWLN GKEYKCKVSN
KALPAPIEKT ISKAKGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP
PVLDSDGSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNH YTQKSLSLSP G(K)

HEK Fab 6B3 WT

MKHLWFFLLL VAAPRWVLSQ VQLQQSGPGL VKPSQTLSLT CAISGDSVSS NSAAWNWIRQ SPGRGLEWLG
RTYYRSKWYN DYAVSVKSRI TINPDTSKNQ FSLQLNSVTP EDTAVYYCAR SYFISFFSFD YWGQGTLVTV
SSASTKGPSV FPLAPSSKST SGGTAALGCL VKDYFPEPVT VSWNSGALTS GVHTFPAVLQ SSGLYSLSSV
VTVPSSSLGT QTYICNVNHK PSNTKVDKRV EPKSCDKTHL EQKLISEEDL NSAVDHHHHH H

HEK Fab 6B3 M

MKHLWFFLLL VAAPRWVLSQ VQLVQSGPGL VKPGQTLSLT CAISGDSVSS NSAAWNWIRQ SPGRGLEWLG
RIYYRSKWYN DYADSVKGRI TINPDTSKNQ FYLQLNSVTP EDTAVYYCAR SYFISFFSFD YWGQGTLVTV
SSASTKGPSV FPLAPSSKST SGGTAALGCL VKDYFPEPVT VSWNSGALTS GVHTFPAVLQ SSGLYSLSSV
VTVPSSSLGT QTYICNVNHK PSNTKVDKRV EPKSCDKTHL EQKLISEEDL NSAVDHHHHH H

HEK IgG 6B3 lambda

MAWALLLLTL LTQGTGSWAD IELTQPPSVS VAPGQTARIS CSGDALGDKY ASWYQQKPGQ APVLVIYDDS
DRPSGIPERF SGSNSGNTAT LTISGTQAED EADYYCQSYD SGFSTVFGGG TKLTVLGQPK AAPSVTLFPP
SSEELQANKA TLVCLISDFY PGAVTVAWKG DSSPVKAGVE TTTPSKQSNN KYAASSYLSL TPEQWKSHRS
YSCQVTHEGS TVEKTVAPTE CS

Pichia (PP) constructs:

PP IgG 2C2 WT

MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YSDLEGDFDV AVLPFSNSTN NGLLFINTTI
ASIAAKEEGV SLEKREAEAA GIQQVQLQQS GPGLVKPSQT LSLTCAISGD SVSSNSAAWN WIRQSPGRGL
EWLGRTYYRS KWYNDYAVSV KSRITINPDT SKNQFSLQLN SVTPEDTAVY YCARQRGHYG KGYKGFNSGF
FDFWGQGTLV TVSSASTKGP SVFPLAPSSK STSGGTAALG CLVKDYFPEP VTVSWNSGAL TSGVHTFPAV
LQSSGLYSLS SVVTVPSSSL GTQTYICNVN HKPSNTKVDK RVEPKSCDKT HTCPPCPAPE LLGGPSVFLF
PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNS*T*YRVV SVLTVLHQDW
LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFY PSDIAVEWES
NGQPENNYKT TPPVLDSDGS FFLYSKLTVD KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGK

PP IgG 2C2 M

MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YSDLEGDFDV AVLPFSNSTN NGLLFINTTI
ASIAAKEEGV SLEKREAEAA GIQQVQLVQS GPGLVKPGQT LSLTCAISGD SVSSNSAAWN WIRQSPGRGL
EWLGRIYYRS KWYNDYADSV KGRITINPDT SKNQFYLQLN SVTPEDTAVY YCARQRGHYG KGYKGFNSGF
FDFWGQGTLV TVSSASTKGP SVFPLAPSSK STSGGTAALG CLVKDYFPEP VTVSWNSGAL TSGVHTFPAV
LQSSGLYSLS SVVTVPSSSL GTQTYICNVN HKPSNTKVDK RVEPKSCDKT HTCPPCPAPE LLGGPSVFLF
PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNS*T*YRVV SVLTVLHQDW
LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFY PSDIAVEWES
NGQPENNYKT TPPVLDSDGS FFLYSKLTVD KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGK

PP Fab 2C2 WT

MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YSDLEGDFDV AVLPFSNSTN NGLLFINTTI
ASIAAKEEGV SLEKREAEAA GIQQVQLQQS GPGLVKPSQT LSLTCAISGD SVSSNSAAWN WIRQSPGRGL
EWLGRTYYRS KWYNDYAVSV KSRITINPDT SKNQFSLQLN SVTPEDTAVY YCARQRGHYG KGYKGFNSGF
FDFWGQGTLV TVSSASTKGP SVFPLAPSSK STSGGTAALG CLVKDYFPEP VTVSWNSGAL TSGVHTFPAV
LQSSGLYSLS SVVTVPSSSL GTQTYICNVN HKPSNTKVDK RVEPKSCDKT HLEQKLISEE DLNSAVDHHH HHH

PP Fab 2C2 WT

MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YSDLEGDFDV AVLPFSNSTN NGLLFINTTI
ASIAAKEEGV SLEKREAEAA GIQQVQLVQS GPGLVKPGQT LSLTCAISGD SVSSNSAAWN WIRQSPGRGL
EWLGRIYYRS KWYNDYADSV KGRITINPDT SKNQFYLQLN SVTPEDTAVY YCARQRGHYG KGYKGFNSGF
FDFWGQGTLV TVSSASTKGP SVFPLAPSSK STSGGTAALG CLVKDYFPEP VTVSWNSGAL TSGVHTFPAV
LQSSGLYSLS SVVTVPSSSL GTQTYICNVN HKPSNTKVDK RVEPKSCDKT HLEQKLISEE DLNSAVDHHH HHH

PP IgG 2C2 kappa

MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YSDLEGDFDV AVLPFSNSTN NGLLFINTTI
ASIAAKEEGV SLEKREAEAA GIQDIVLTQS PATLSLSPGE RATLSCRASQ SVSSSYLAWY QQKPGQAPRL
LIYGASSRAT GVPARFSGSG SGTDFTLTIS SLEPEDFATY YCQQYYNIPT FGQGTKVEIK RTVAAPSVFI
FPPSDEQLKS GTASVVCLLN NFYPREAKVQ WKVDNALQSG NSQESVTEQD SKDSTYSLSS TLTLSKADYE
KHKVYACEVT HQGLSSPVTK SFNRGEC

PP IgG 6B3 WT

MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YSDLEGDFDV AVLPFSNSTN NGLLFINTTI
ASIAAKEEGV SLEKR**EAEAA** **GIQ**QVQL**Q**QS GPGLVKP**S**QT LSLTCAISGD SVSSNSAAWN WIRQSPGRGL
EWLGR**T**YYRS KWYNDYA**V**SV K**S**RITINPDT SKNQF**S**LQLN SVTPEDTAVY YCARSYFISF FSFDYWGQGT
LVTVSSASTK GPSVFPLAPS SKSTSGGTAA LGCLVKDYFP EPVTVSWNSG ALTSGVHTFP AVLQSSGLYS
LSSVVTVPSS SLGTQTYICN VNHKPSNTKV DKRVEPKSCD KTHTCPPCPA PELLGGPSVF LFPPKPKDTL
MISRTPEVTC VVVDVSHEDP EVKFNWYVDG VEVHNAKTKP REEQYNS*T*YR VVSVLTVLHQ DWLNGKEYKC
KVSNKALPAP IEKTISKAKG QPREPQVYTL PPSREEMTKN QVSLTCLVKG FYPSDIAVEW ESNGQPENNY
KTTPPVLDSD GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA LHNHYTQKSL SLSPGK

PP IgG 6B3 M

MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YSDLEGDFDV AVLPFSNSTN NGLLFINTTI
ASIAAKEEGV SLEKR**EAEAA** **GIQ**QVQL**V**QS GPGLVKP**G**QT LSLTCAISGD SVSSNSAAWN WIRQSPGRGL
EWLGR**I**YYRS KWYNDYA**D**SV K**G**RITINPDT SKNQF**Y**LQLN SVTPEDTAVY YCARSYFISF FSFDYWGQGT
LVTVSSASTK GPSVFPLAPS SKSTSGGTAA LGCLVKDYFP EPVTVSWNSG ALTSGVHTFP AVLQSSGLYS
LSSVVTVPSS SLGTQTYICN VNHKPSNTKV DKRVEPKSCD KTHTCPPCPA PELLGGPSVF LFPPKPKDTL
MISRTPEVTC VVVDVSHEDP EVKFNWYVDG VEVHNAKTKP REEQYNSTYR VVSVLTVLHQ DWLNGKEYKC
KVSNKALPAP IEKTISKAKG QPREPQVYTL PPSREEMTKN QVSLTCLVKG FYPSDIAVEW ESNGQPENNY
KTTPPVLDSD GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA LHNHYTQKSL SLSPGK

PP Fab 6B3 WT

MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YSDLEGDFDV AVLPFSNSTN NGLLFINTTI
ASIAAKEEGV SLEKR**EAEAA** **GIQ**QVQL**Q**QS GPGLVKP**S**QT LSLTCAISGD SVSSNSAAWN WIRQSPGRGL
EWLGR**T**YYRS KWYNDYA**V**SV K**S**RITINPDT SKNQF**S**LQLN SVTPEDTAVY YCARSYFISF FSFDYWGQGT
LVTVSSASTK GPSVFPLAPS SKSTSGGTAA LGCLVKDYFP EPVTVSWNSG ALTSGVHTFP AVLQSSGLYS
LSSVVTVPSS SLGTQTYICN VNHKPSNTKV DKRVEPKSCD KTHLEQKLIS EEDLNSAVDH HHHHH

PP Fab 6B3 M

MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YSDLEGDFDV AVLPFSNSTN NGLLFINTTI
ASIAAKEEGV SLEKR**EAEAA** **GIQ**QVQL**V**QS GPGLVKP**G**QT LSLTCAISGD SVSSNSAAWN WIRQSPGRGL
EWLGR**I**YYRS KWYNDYA**D**SV K**G**RITINPDT SKNQF**Y**LQLN SVTPEDTAVY YCARSYFISF FSFDYWGQGT
LVTVSSASTK GPSVFPLAPS SKSTSGGTAA LGCLVKDYFP EPVTVSWNSG ALTSGVHTFP AVLQSSGLYS
LSSVVTVPSS SLGTQTYICN VNHKPSNTKV DKRVEPKSCD KTHLEQKLIS EEDLNSAVDH HHHHH

PP IgG 6B3 lambda

MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YSDLEGDFDV AVLPFSNSTN NGLLFINTTI
ASIAAKEEGV SLEKR**EAEAA** **GIQ**DIELTQP PSVSVAPGQT ARISCSGDAL GDKYASWYQQ KPGQAPVLVI
YDDSDRPSGI PERFSGSNSG NTATLTISGT QAEDEADYYC QSYDSGFSTV FGGGTKLTVL GQPKAAPSVT
LFPPSSEELQ ANKATLVCLI SDFYPGAVTV AWKGDSSPVK AGVETTTPSK QSNNKYAASS YLSLTPEQWK
SHRSYSCQVT HEGSTVEKTV APTECS

Prokaryotic constructs (*E. coli*; EC):

EC IgG 2C2 WT/kappa

MKKNIAFLLA SMFVFSIATN AYADIVLTQS PATLSLSPGE RATLSCRASQ SVSSSYLAWY QQKPGQAPRL
LIYGASSRAT GVPARFSGSG SGTDFTLTIS SLEPEDFATY YCQQYYNIPT FGQGTKVEIK RTVAAPSVFI
FPPSDEQLKS GTASVVCLLN NFYPREAKVQ WKVDNALQSG NSQESVTEQD SKDSTYSLSS TLTLSKADYE
KHKVYACEVT HQGLSSPVTK SFNRGEC**A CVGENKMKQS TIALALLPLL FTPVTKAQVQ LQQSGPGLVK
PSQTLSLTCA ISGDSVSSNS AAWNWIRQSP GRGLEWLGRT YYRSKWYNDY AVSVKSRITI NPDTSKNQFS
LQLNSVTPED TAVYYCARQR GHYGKGYKGF NSGFFDFWGQ GTLVTVSSAS TKGPSVFPLA PSSKSTSGGT
AALGCLVKDY FPEPVTVSWN SGALTSGVHT FPAVLQSSGL YSLSSVVTVP SSSLGTQTYI CNVNHKPSNT
KVDKRVEPKS CDKTHTCPPC PAPELLGGPS VFLFPPKPKD TLMISRTPEV TCVVVDVSHE DPEVKFNWYV
DGVEVHNAKT KPREEQYNST YRVVSVLTVL HQDWLNGKEY KCKVSNKALP APIEKTISKA KGQPREPQVY
TLPPSREEMT KNQVSLTCLV KGFYPSDIAV EWESNGQPEN NYKTTPPVLD SDGSFFLYSK LTVDKSRWQQ
GNVFSCSVMH EALHNHYTQK SLSLSPGK

EC IgG 2C2 M/kappa

MKKNIAFLLA SMFVFSIATN AYADIVLTQS PATLSLSPGE RATLSCRASQ SVSSSYLAWY QQKPGQAPRL
LIYGASSRAT GVPARFSGSG SGTDFTLTIS SLEPEDFATY YCQQYYNIPT FGQGTKVEIK RTVAAPSVFI
FPPSDEQLKS GTASVVCLLN NFYPREAKVQ WKVDNALQSG NSQESVTEQD SKDSTYSLSS TLTLSKADYE
KHKVYACEVT HQGLSSPVTK SFNRGEC**A CVGENKMKQS TIALALLPLL FTPVTKAQVQ LVQSGPGLVK
PGQTLSLTCA ISGDSVSSNS AAWNWIRQSP GRGLEWLGRI YYRSKWYNDY ADSVKGRITI NPDTSKNQFY
LQLNSVTPED TAVYYCARQR GHYGKGYKGF NSGFFDFWGQ GTLVTVSSAS TKGPSVFPLA PSSKSTSGGT
AALGCLVKDY FPEPVTVSWN SGALTSGVHT FPAVLQSSGL YSLSSVVTVP SSSLGTQTYI CNVNHKPSNT
KVDKRVEPKS CDKTHTCPPC PAPELLGGPS VFLFPPKPKD TLMISRTPEV TCVVVDVSHE DPEVKFNWYV
DGVEVHNAKT KPREEQYNST YRVVSVLTVL HQDWLNGKEY KCKVSNKALP APIEKTISKA KGQPREPQVY
TLPPSREEMT KNQVSLTCLV KGFYPSDIAV EWESNGQPEN NYKTTPPVLD SDGSFFLYSK LTVDKSRWQQ
GNVFSCSVMH EALHNHYTQK SLSLSPGK

EC Fab 2C2 WT/kappa

MKKTAIAIAV ALAGFATVAQ ADIVLTQSPA TLSLSPGERA TLSCRASQSV SSSYLAWYQQ KPGQAPRLLI
YGASSRATGV PARFSGSGSG TDFTLTISSL EPEDFATYYC QQYYNIPTFG QGTKVEIKRT VAAPSVFIFP
PSDEQLKSGT ASVVCLLNNF YPREAKVQWK VDNALQSGNS QESVTEQDSK DSTYSLSSTL TLSKADYEKH
KVYACEVTHQ GLSSPVTKSF NRGEC**ACV GENKMKQSTI ALALLPLLFT PVTKAQVQLQ QSGPGLVKPS
QTLSLTCAIS GDSVSSNSAA WNWIRQSPGR GLEWLGRTYY RSKWYNDYAV SVKSRITINP DTSKNQFSLQ
LNSVTPEDTA VYYCARQRGH YGKGYKGFNS GFFDFWGQGT LVTVSSASTK GPSVFPLAPS SKSTSGGTAA
LGCLVKDYFP EPVTVSWNSG ALTSGVHTFP AVLQSSGLYS LSSVVTVPSS SLGTQTYICN VNHKPSNTKV
DKRVEPKSCD KTHLEQKLIS EEDLNSAVDH HHHHH

EC Fab 2C2 M/kappa

MKKTAIAIAV ALAGFATVAQ ADIVLTQSPA TLSLSPGERA TLSCRASQSV SSSYLAWYQQ KPGQAPRLLI
YGASSRATGV PARFSGSGSG TDFTLTISSL EPEDFATYYC QQYYNIPTFG QGTKVEIKRT VAAPSVFIFP
PSDEQLKSGT ASVVCLLNNF YPREAKVQWK VDNALQSGNS QESVTEQDSK DSTYSLSSTL TLSKADYEKH
KVYACEVTHQ GLSSPVTKSF NRGEC**ACV GENKMKQSTI ALALLPLLFT PVTKAQVQLV QSGPGLVKPG
QTLSLTCAIS GDSVSSNSAA WNWIRQSPGR GLEWLGRIYY RSKWYNDYAD SVKGRITINP DTSKNQFYLQ
LNSVTPEDTA VYYCARQRGH YGKGYKGFNS GFFDFWGQGT LVTVSSASTK GPSVFPLAPS SKSTSGGTAA
LGCLVKDYFP EPVTVSWNSG ALTSGVHTFP AVLQSSGLYS LSSVVTVPSS SLGTQTYICN VNHKPSNTKV
DKRVEPKSCD KTHLEQKLIS EEDLNSAVDH HHHHH

EC IgG 6B3 WT/lambda

MKKNIAFLLA SMFVFSIATN AYADIELTQP PSVSVAPGQT ARISCSGDAL GDKYASWYQQ KPGQAPVLVI
YDDSDRPSGI PERFSGSNSG NTATLTISGT QAEDEADYYC QSYDSGFSTV FGGGTKLTVL GQPKAAPSVT
LFPPSSEELQ ANKATLVCLI SDFYPGAVTV AWKGDSSPVK AGVETTTPSK QSNNKYAASS YLSLTPEQWK
SHRSYSCQVT HEGSTVEKTV APTECS**AC VGENKMKQST IALALLPLLF TPVTKAQVQL QQSGPGLVKP
SQTLSLTCAI SGDSVSSNSA AWNWIRQSPG RGLEWLGRTY YRSKWYNDYA VSVKSRITIN PDTSKNQFSL
QLNSVTPEDT AVYYCARSYF ISFFSFDYWG QGTLVTVSSA STKGPSVFPL APSSKSTSGG TAALGCLVKD
YFPEPVTVSW NSGALTSGVH TFPAVLQSSG LYSLSSVVTV PSSSLGTQTY ICNVNHKPSN TKVDKRVEPK
SCDKTHTCPP CPAPELLGGP SVFLFPPKPK DTLMISRTPE VTCVVVDVSH EDPEVKFNWY VDGVEVHNAK
TKPREEQYNS TYRVVSVLTV LHQDWLNGKE YKCKVSNKAL PAPIEKTISK AKGQPREPQV YTLPPSREEM
TKNQVSLTCL VKGFYPSDIA VEWESNGQPE NNYKTTPPVL DSDGSFFLYS KLTVDKSRWQ QGNVFSCSVM
HEALHNHYTQ KSLSLSPGK

EC IgG 6B3 M/lambda

MKKNIAFLLA SMFVFSIATN AYADIELTQP PSVSVAPGQT ARISCSGDAL GDKYASWYQQ KPGQAPVLVI
YDDSDRPSGI PERFSGSNSG NTATLTISGT QAEDEADYYC QSYDSGFSTV FGGGTKLTVL GQPKAAPSVT
LFPPSSEELQ ANKATLVCLI SDFYPGAVTV AWKGDSSPVK AGVETTTPSK QSNNKYAASS YLSLTPEQWK
SHRSYSCQVT HEGSTVEKTV APTECS**AC VGENKMKQST IALALLPLLF TPVTKAQVQL VQSGPGLVKP
GQTLSLTCAI SGDSVSSNSA AWNWIRQSPG RGLEWLGRIY YRSKWYNDYA DSVKGRITIN PDTSKNQFYL
QLNSVTPEDT AVYYCARSYF ISFFSFDYWG QGTLVTVSSA STKGPSVFPL APSSKSTSGG TAALGCLVKD
YFPEPVTVSW NSGALTSGVH TFPAVLQSSG LYSLSSVVTV PSSSLGTQTY ICNVNHKPSN TKVDKRVEPK
SCDKTHTCPP CPAPELLGGP SVFLFPPKPK DTLMISRTPE VTCVVVDVSH EDPEVKFNWY VDGVEVHNAK
TKPREEQYNS TYRVVSVLTV LHQDWLNGKE YKCKVSNKAL PAPIEKTISK AKGQPREPQV YTLPPSREEM
TKNQVSLTCL VKGFYPSDIA VEWESNGQPE NNYKTTPPVL DSDGSFFLYS KLTVDKSRWQ QGNVFSCSVM
HEALHNHYTQ KSLSLSPGK

EC Fab 6B3 WT/lambda

MKKTAIAIAV ALAGFATVAQ ADIELTQPPS VSVAPGQTAR ISCSGDALGD KYASWYQQKP GQAPVLVIYD
DSDRPSGIPE RFSGSNSGNT ATLTISGTQA EDEADYYCQS YDSGFSTVFG GGTKLTVLGQ PKAAPSVTLF
PPSSEELQAN KATLVCLISD FYPGAVTVAW KGDSSPVKAG VETTTPSKQS NNKYAASSYL SLTPEQWKSH
RSYSCQVTHE GSTVEKTVAP TECS**ACVG ENKMKQSTIA LALLPLLFTP VTKAQVQLQQ SGPGLVKPSQ
TLSLTCAISG DSVSSNSAAW NWIRQSPGRG LEWLGRTYYR SKWYNDYAVS VKSRITINPD TSKNQFSLQL
NSVTPEDTAV YYCARSYFIS FFSFDYWGQG TLVTVSSAST KGPSVFPLAP SSKSTSGGTA ALGCLVKDYF
PEPVTVSWNS GALTSGVHTF PAVLQSSGLY SLSSVVTVPS SSLGTQTYIC NVNHKPSNTK VDKRVEPKSC
DKTHLEQKLI SEEDLNSAVD HHHHHH

EC Fab 6B3 M/lambda

MKKTAIAIAV ALAGFATVAQ ADIELTQPPS VSVAPGQTAR ISCSGDALGD KYASWYQQKP GQAPVLVIYD
DSDRPSGIPE RFSGSNSGNT ATLTISGTQA EDEADYYCQS YDSGFSTVFG GGTKLTVLGQ PKAAPSVTLF
PPSSEELQAN KATLVCLISD FYPGAVTVAW KGDSSPVKAG VETTTPSKQS NNKYAASSYL SLTPEQWKSH
RSYSCQVTHE GSTVEKTVAP TECS**ACVG ENKMKQSTIA LALLPLLFTP VTKAQVQLVQ SGPGLVKPGQ
TLSLTCAISG DSVSSNSAAW NWIRQSPGRG LEWLGRIYYR SKWYNDYADS VKGRITINPD TSKNQFYLQL
NSVTPEDTAV YYCARSYFIS FFSFDYWGQG TLVTVSSAST KGPSVFPLAP SSKSTSGGTA ALGCLVKDYF
PEPVTVSWNS GALTSGVHTF PAVLQSSGLY SLSSVVTVPS SSLGTQTYIC NVNHKPSNTK VDKRVEPKSC
DKTHLEQKLI SEEDLNSAVD HHHHHH

5.4 Conferences, Posters and Oral presentations

Pichia 2012: oral and poster presentation

"Improving the biophysical properties of full-length IgGs by using *P. pastoris* as expression host"
February 29th - March 3rd, 2012 in Alpbach / Austria

3rd PEGS Europe - Protein and Antibody Engineering Summit: oral presentation

"Folding and Aggregation of IgGs - Influence of Mutations & Expression Systems"
October 11th - 13th, 2011 in Hannover / Germany

BioProcessing, Biologics & Biotherapeutics: oral presentation

"Folding and Aggregation of IgGs - Influence of Mutations & Expression Systems"
July 20th - 21st, 2011 in Edinburgh / Scotland

4th Halle Conference - Recombinant Protein Production: oral presentation

"Folding and Aggregation of IgGs - Influence of Mutations & Expression Systems"
February 24th - 26th, 2011 in Halle / Germany

IBC's 21st Annual Antibody Engineering & Antibody Therapeutics: poster presentation

"Folding and Aggregation in the IgG Format: Influence of Mutations and the Expression System"
December 5th - 9th, 2010 in San Diego / USA

VIII European Symposium of the Protein Society: poster presentation

"Stability engineering of antibody variable domains improves their biophysical properties in the context of IgGs produced in eukaryotic cells"
June 14th - 18th, 2009 in Zurich / Switzerland

7th International Conference on Recombinant Antibodies

June 23rd - 27th, 2008 in Dublin / Ireland

Antibody Europe: The Next Generation of Antibodies

November 7th - 8th, 2007 in Vienna / Austria

5.5 Curriculum Vitae

Name:	Jonas V. Schaefer
Date of birth:	July 30th 1980
Place of birth:	Heidelberg / Germany
Nationality:	German

EDUCATION

03 / 2006 – 03 / 2012	**PhD Studies in the international 'Molecular Life Science' (MLS) Program of the ETH and University of Zurich / Switzerland** Field of Research: 'Protein Design and Antibody Engineering' in the laboratory of Prof. A. Plückthun (Department of Biochemistry) Title of Thesis: "Folding and Aggregation of IgGs – Influence of Mutations and Expression Systems"
09 / 2004 – 01 / 2006	**Master Studies in 'Molecular and Cellular Biology' at the University of Massachusetts Amherst / USA** Field of Study: 'Viral Propagation and Cellular Biochemistry' in the laboratory of Prof. D. Hebert (Department of Biochemistry and Molecular Biology) Title of Thesis: "Small T-antigen enhances SV40 propagation by decreasing the duration of the viral life cycle though an increase in late gene production" Master´s Degree with a GPA of 4.0 (scale 4.0 – 1.0; 4.0 being best)
10 / 2001 – 07 / 2004	**Undergraduate Studies in 'Molecular Biotechnology' at the Ruprecht-Karls-University Heidelberg / Germany** Major: 'Drug Design', Minors: Biophysics and Bioinformatics Bachelor´s Degree with grade average of 1.1 (scale 1.0 – 6.0; 1.0 being best)
09 / 1991 – 06 / 2000	**High School St. Raphael Gymnasium, Heidelberg / Germany** Majors: Mathematics and Physics A-Levels with final grade of 1.0 (scale 1.0 – 6.0; 1.0 being best)

SCHOLARSHIPS

12 / 2010 – 10 / 2011	**Travel grants for the attendances of various conferences** granted by the University of Zurich, the Life Science Zurich Graduate School, the Hartmann Müller Foundation and the Swiss Society of Biochemistry
11 / 2006 – 11 / 2008	**Kekulé-Scholar of the 'German Chemical Industry Association'** Two year grant for PhD Studies
07 / 2005 – 01 / 2006	**Fellowship granted by the 'Landesstiftung Baden-Württemberg'** Support of Master Studies in the USA
09 / 2004 – 06 / 2005	**Scholarship by the 'Fulbright Foundation'** for studying abroad in the USA
11 / 2001 – 01 / 2006	**Scholar of the 'National German Academic Foundation'**